Programme

Time / Part	Title	Description and content
Part 1: 9:00 – 12:30	**Sentinel-2 and spatial image analysis**	1 – Sentinel-2 introduction Exercise 1: „*Introduction into Sentinel-2*“ 2 – Spatial image analysis
Part 2: 13:30 – 15:00	**Technologies for big EO image analysis**	3 – Mission planning with the EO-Compass Exercise 2: „*Mission planning with the EO-Compass*“ 4 – Big Earth observation data analytics - overview
Part 3: 15:15 – 17:30	**Concepts for time series analysis for land use / land cover**	5 – Temporal dimension of EO data 6 – Time series analysis of EO data Exercise 3: „*Investigate forest fires in Andalusia, Spain using a time series of Sentinel-2 images*“

3

Mission planning with the EO-Compass

Concepts and technologies for big EO analytics

TAT-7 | 21.06.2019 | Martin Sudmanns, Dirk Tiede, & contrib. Z_GIS colleagues

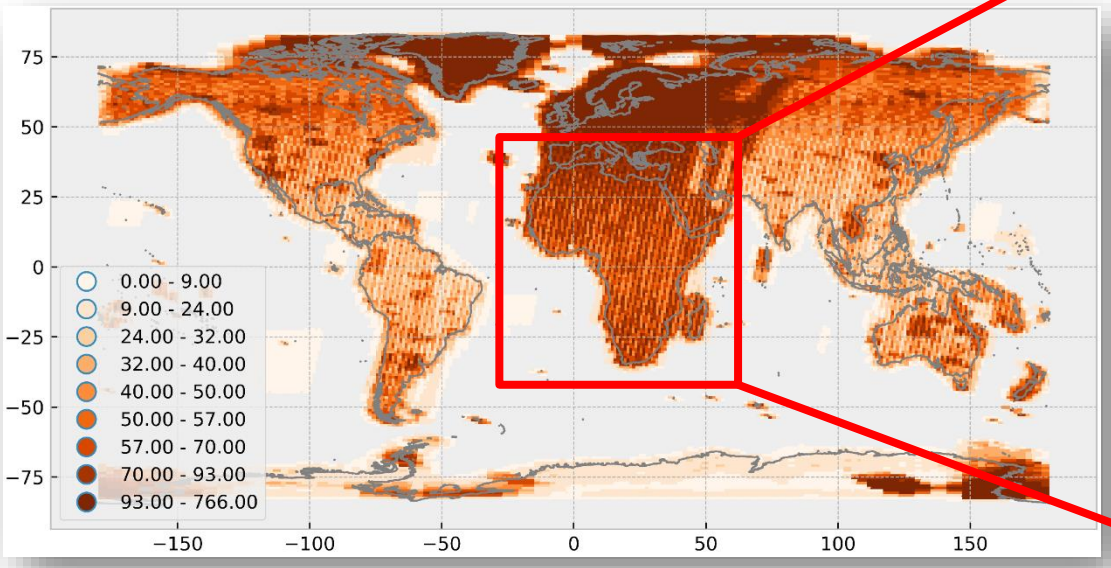
Interfaculty Department of Geoinformatics – Z_GIS, University of Salzburg

EO-Compass

The **data coverage** is not homogeneous and influences the success of analyses

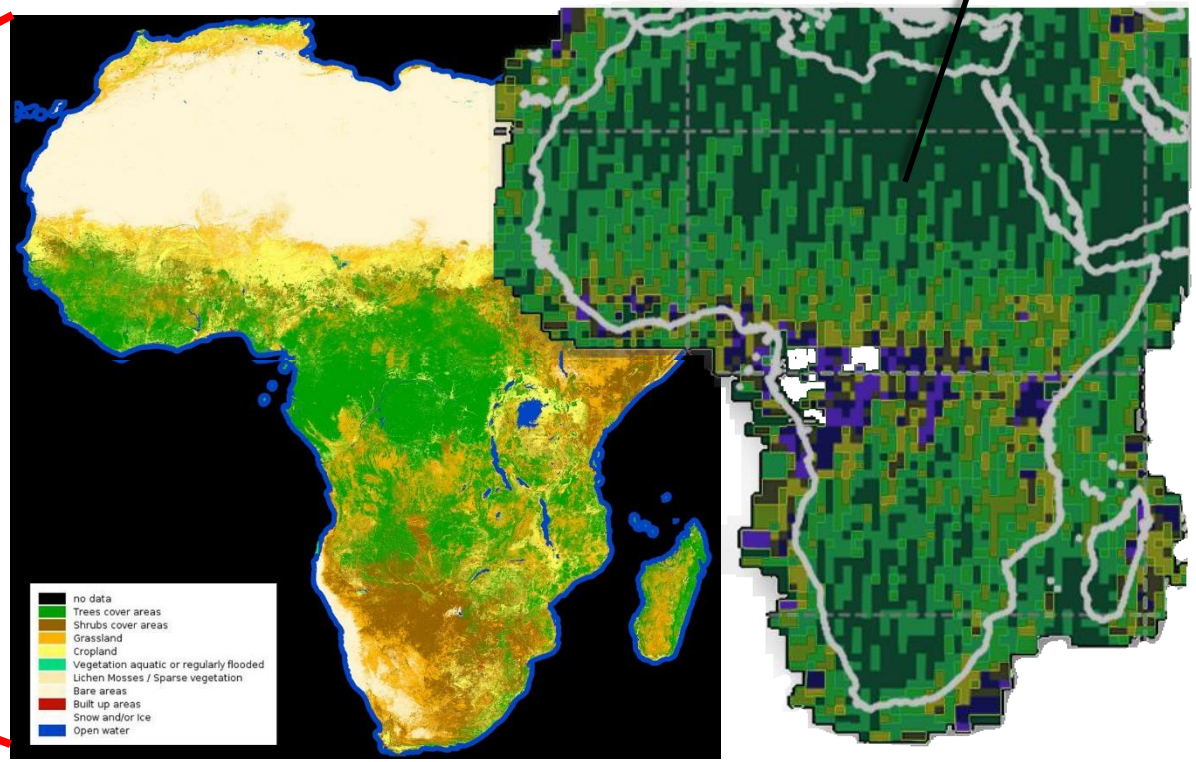
Decisions on feasibility is **difficult** and requires lots of **expert-knowledge**

Number of acquired Sentinel-2 A scenes



Classification result

Suitability for time series analysis, green is high



EO-Compass

Users **consult** the EO-Compass to **systematically evaluate** the feasibility of their ideas.

Key questions

How to know whether an idea is feasible?

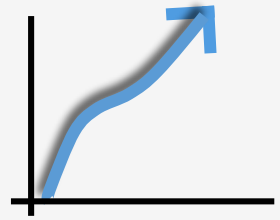
Can we estimate it before spending time and money?

>6 mil.

Sentinel-2 Scenes in the archive

3.4 TB

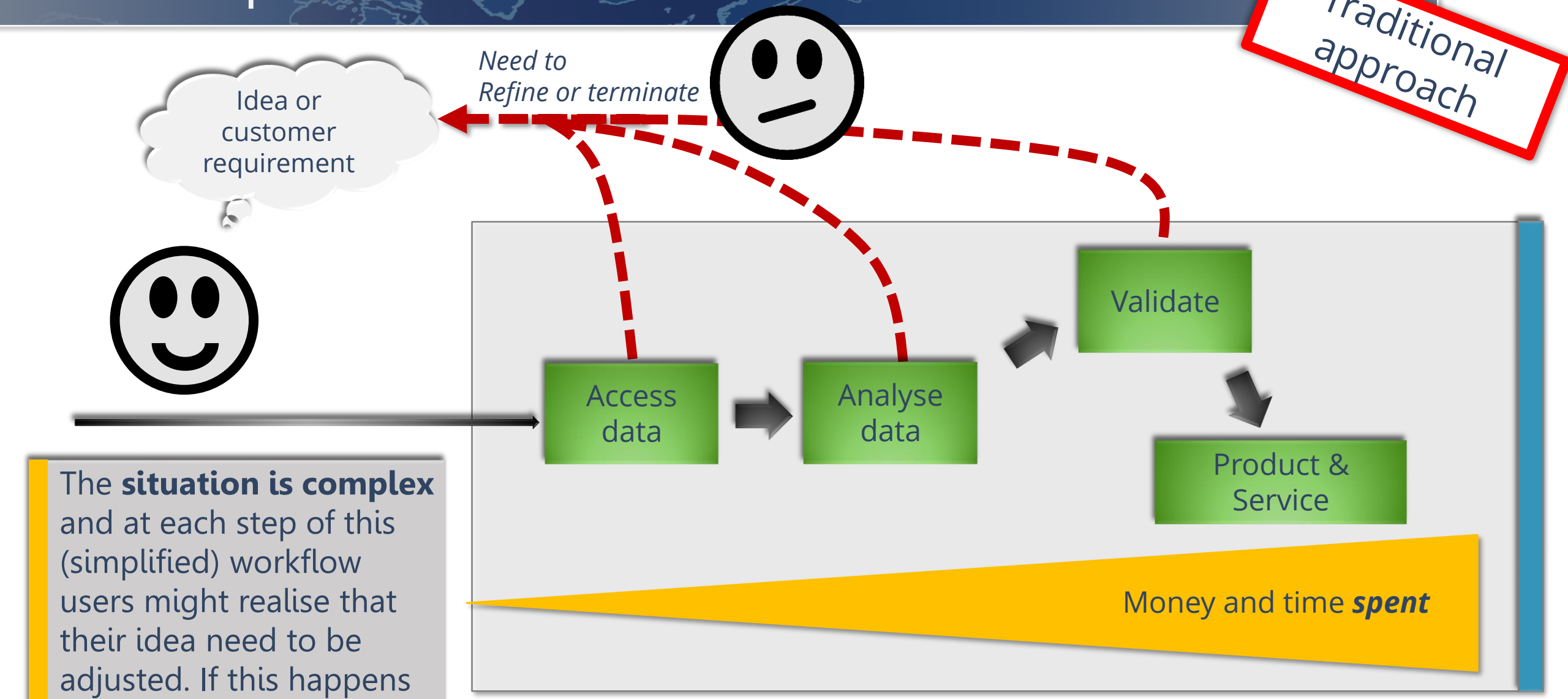
new data volume every day



70 %

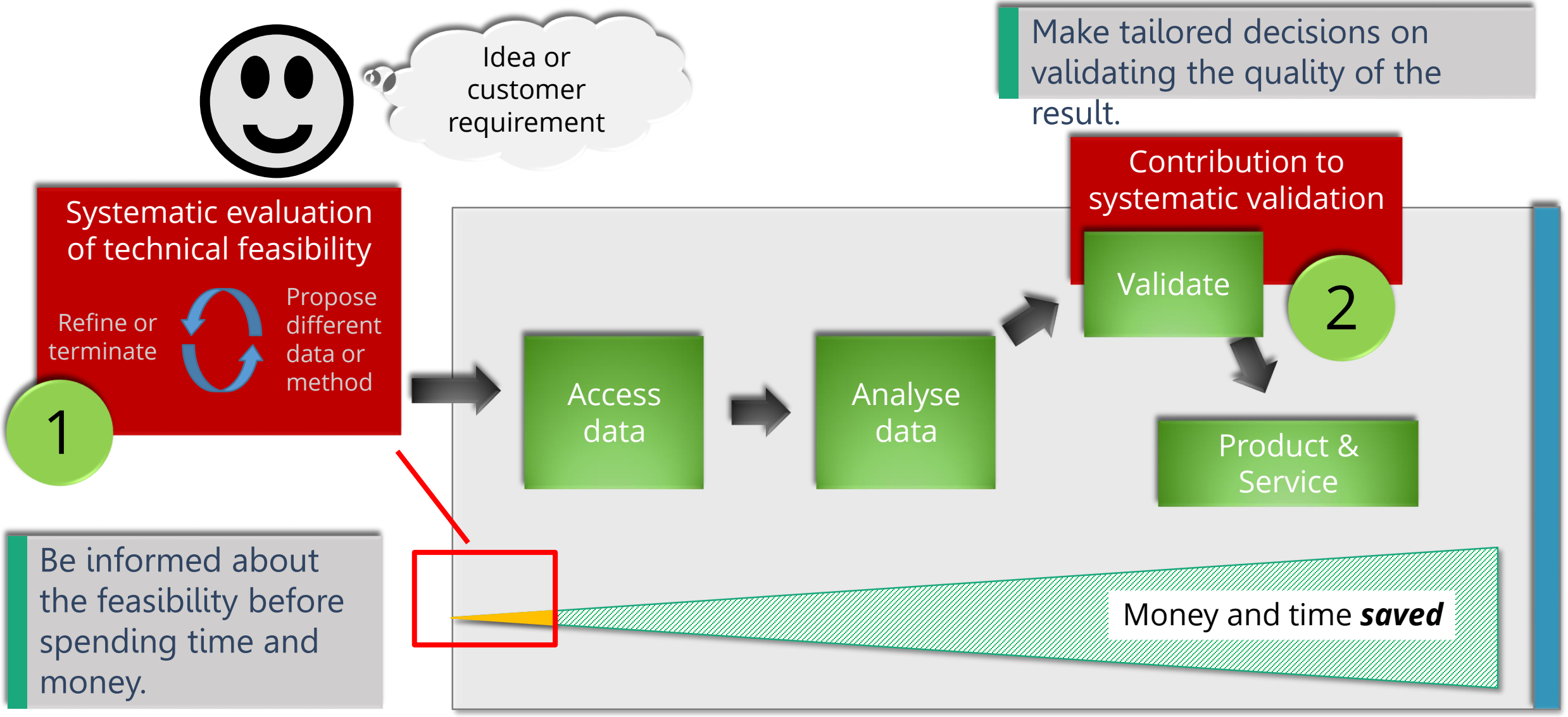
of the Earth is cloud covered any time

EO-Compass

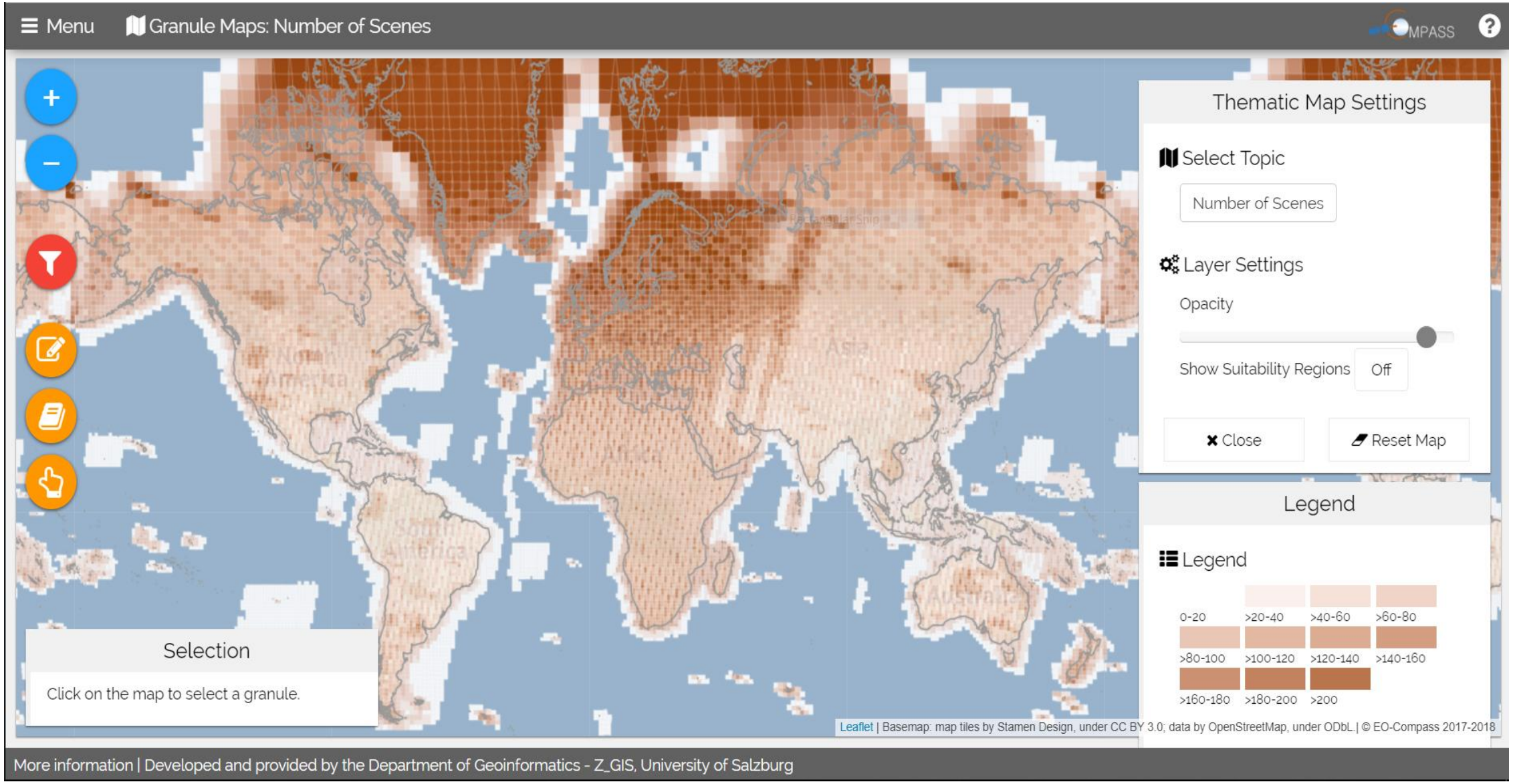


The **situation is complex** and at each step of this (simplified) workflow users might realise that their idea need to be adjusted. If this happens not early enough, **time and money is lost**.

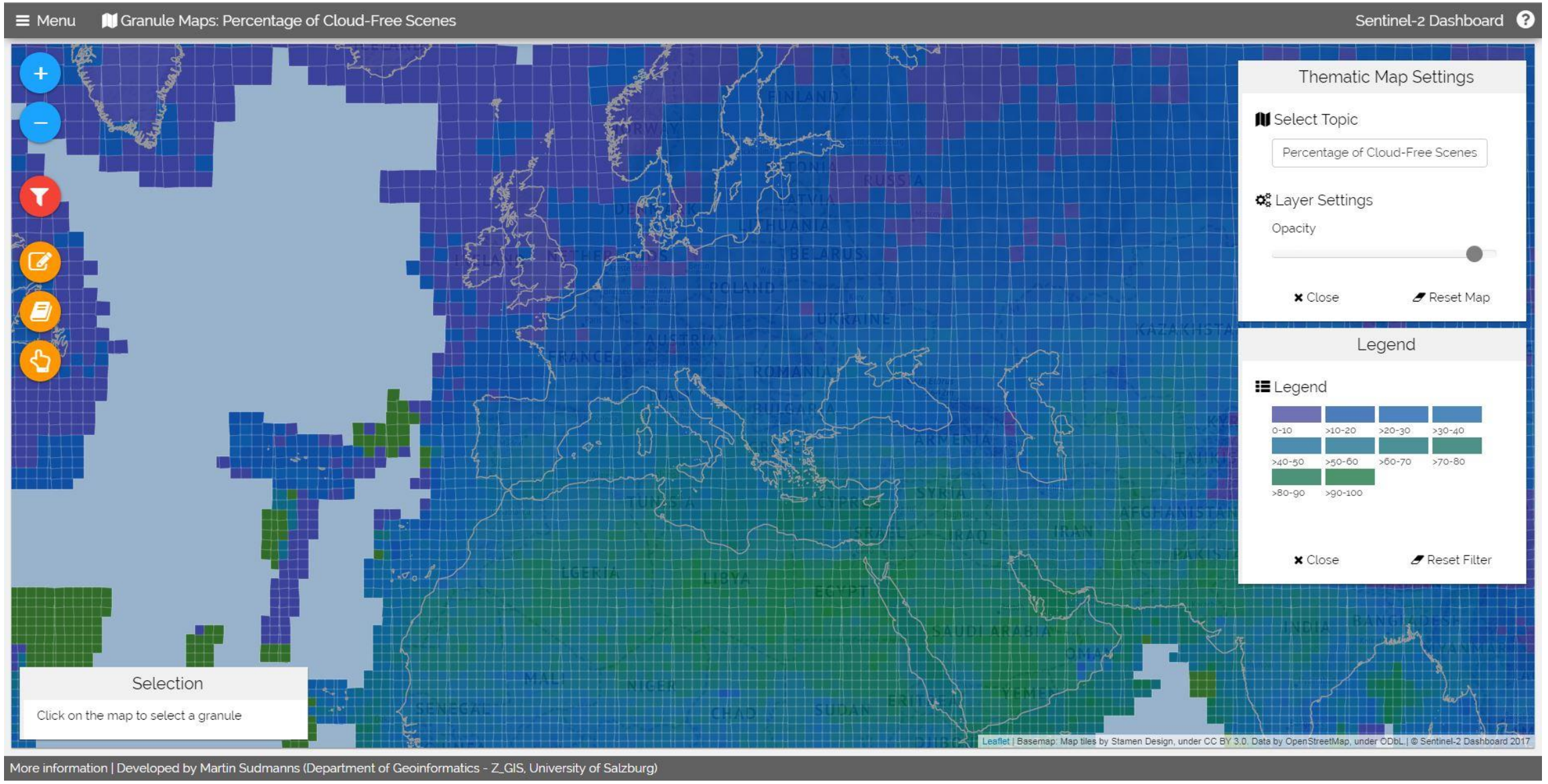
EO-Compass


Idea or customer requirement
Make tailored decisions on validating the quality of the result.
Systematic evaluation of technical feasibility
Refine or terminate
Propose different data or method
1
Access data
Analyse data
Contribution to systematic validation
Validate
2
Product & Service
Be informed about the feasibility before spending time and money.
Money and time saved

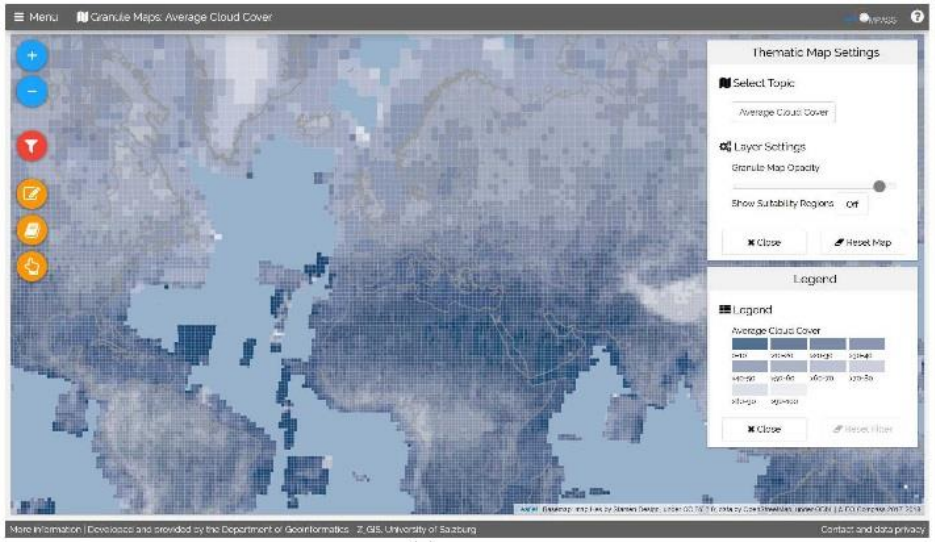
EO-Compass


Menu
Granule Maps: Number of Scenes
MPASS
Thematic Map Settings
Select Topic
Number of Scenes
Layer Settings
Opacity
Show Suitability Regions
Off
Close
Reset Map
Legend
Legend
0-20
>20-40
>40-60
>60-80
>80-100
>100-120
>120-140
>140-160
>160-180
>180-200
>200
Selection
Click on the map to select a granule.
Leaflet | Basemap: map tiles by Stamen Design, under CC BY 3.0; data by OpenStreetMap, under ODbL.| © EO-Compass 2017-2018
More information | Developed and provided by the Department of Geoinformatics - Z_GIS, University of Salzburg

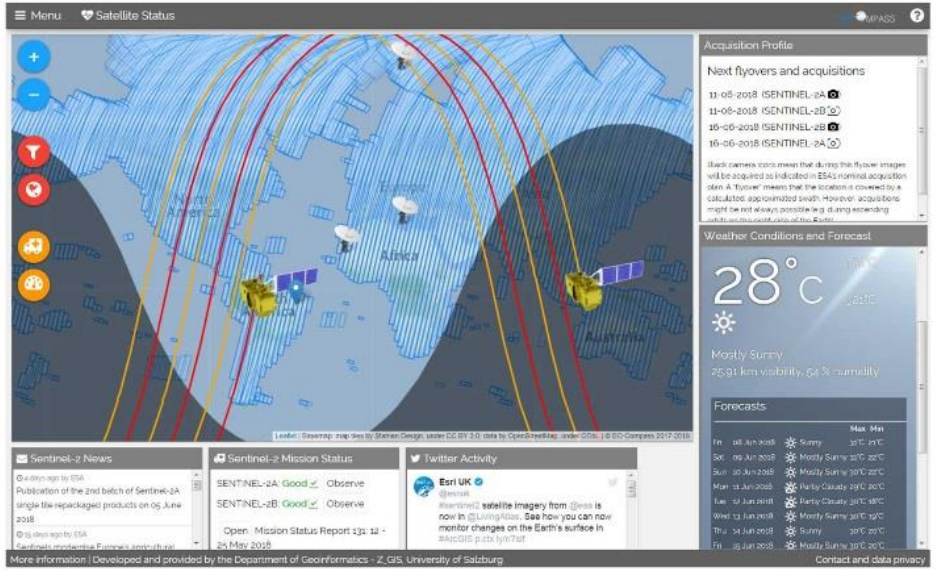
EO-Compass


Menu
Granule Maps: Percentage of Cloud-Free Scenes
Sentinel-2 Dashboard
Thematic Map Settings
Select Topic
Percentage of Cloud-Free Scenes
Layer Settings
Opacity
Close
Reset Map
Legend
Legend
0-10
>10-20
>20-30
>30-40
>40-50
>50-60
>60-70
>70-80
>80-90
>90-100
Close
Reset Filter
Selection
Click on the map to select a granule
Leaflet | Basemap: Map tiles by Stamen Design, under CC BY 3.0. Data by OpenStreetMap, under ODbL. | © Sentinel-2 Dashboard 2017
More information | Developed by Martin Sudmanns (Department of Geoinformatics - Z_GIS, University of Salzburg)

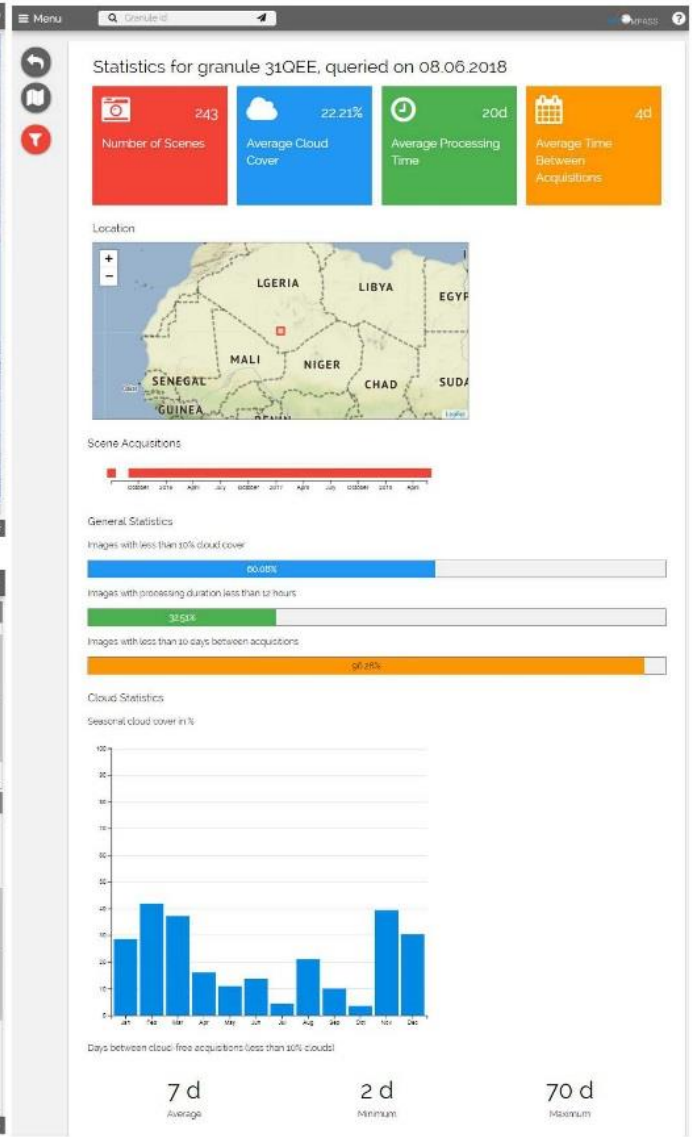
EO-Compass



(a)

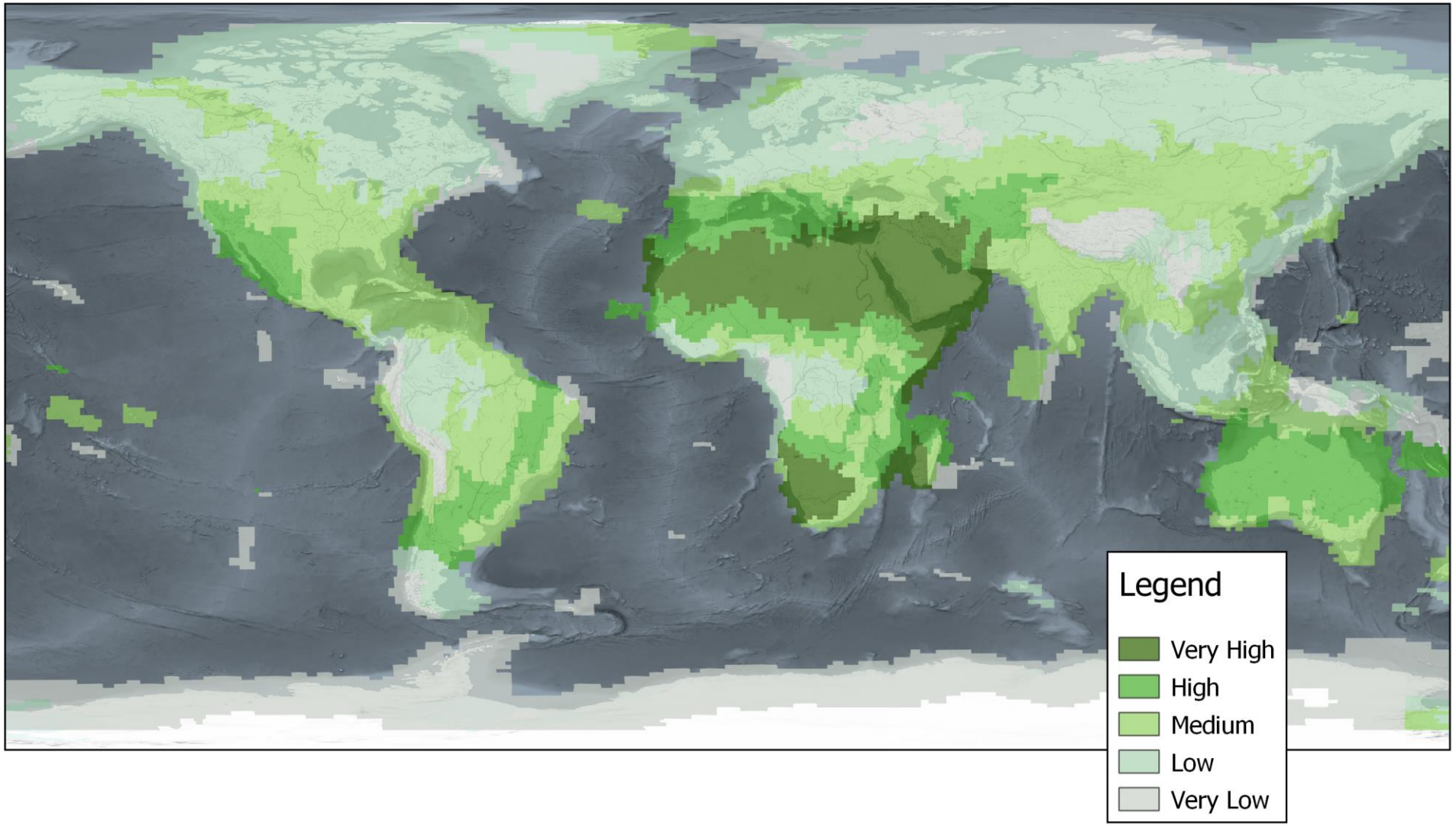


(b)



(c)

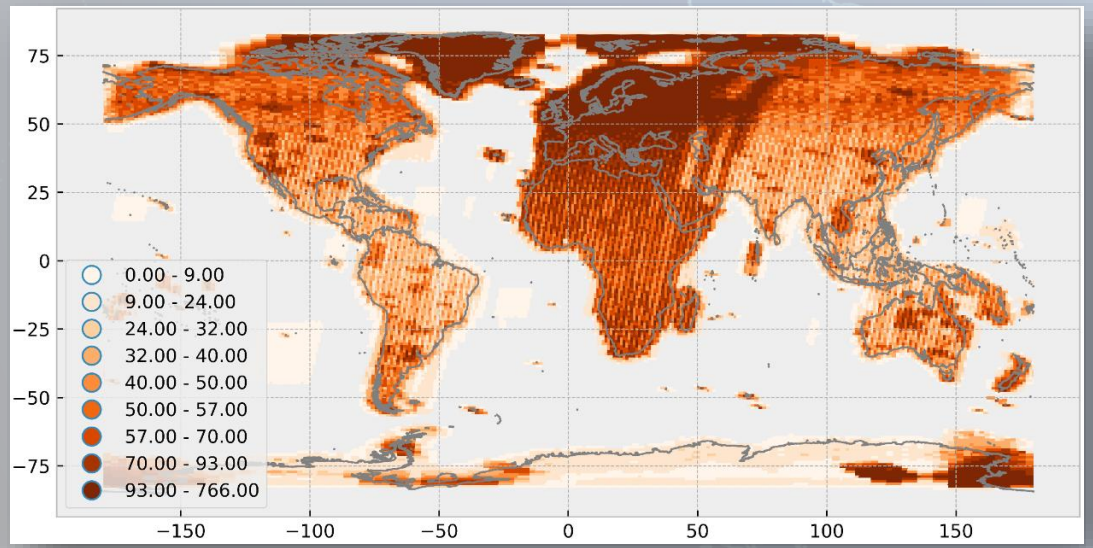
EO-Compass

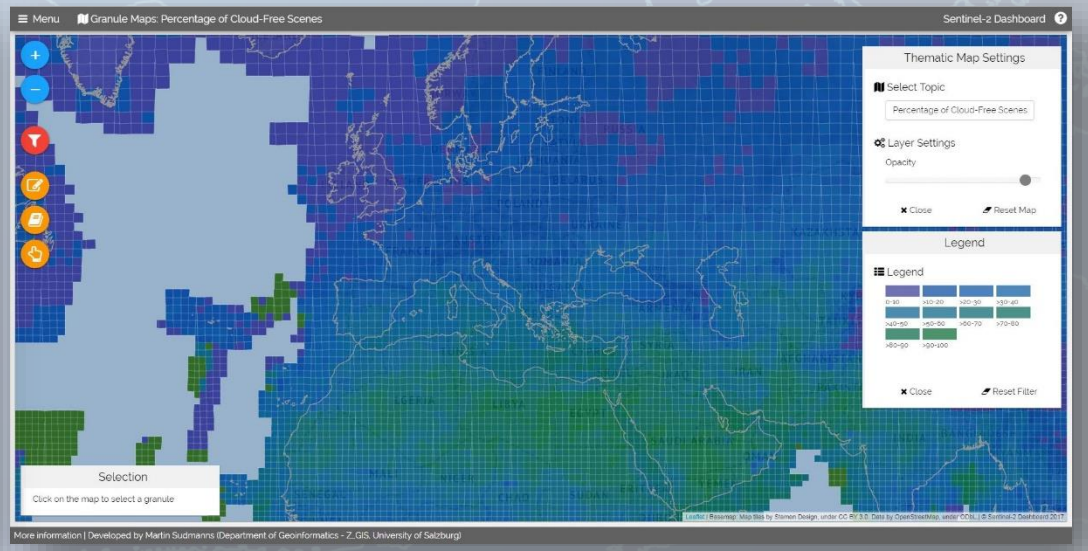


Exercise 2:

Note: The acquisitions are currently uncomplete as there is a major update pending …

Investigate global availability of Sentinel-2 images





4

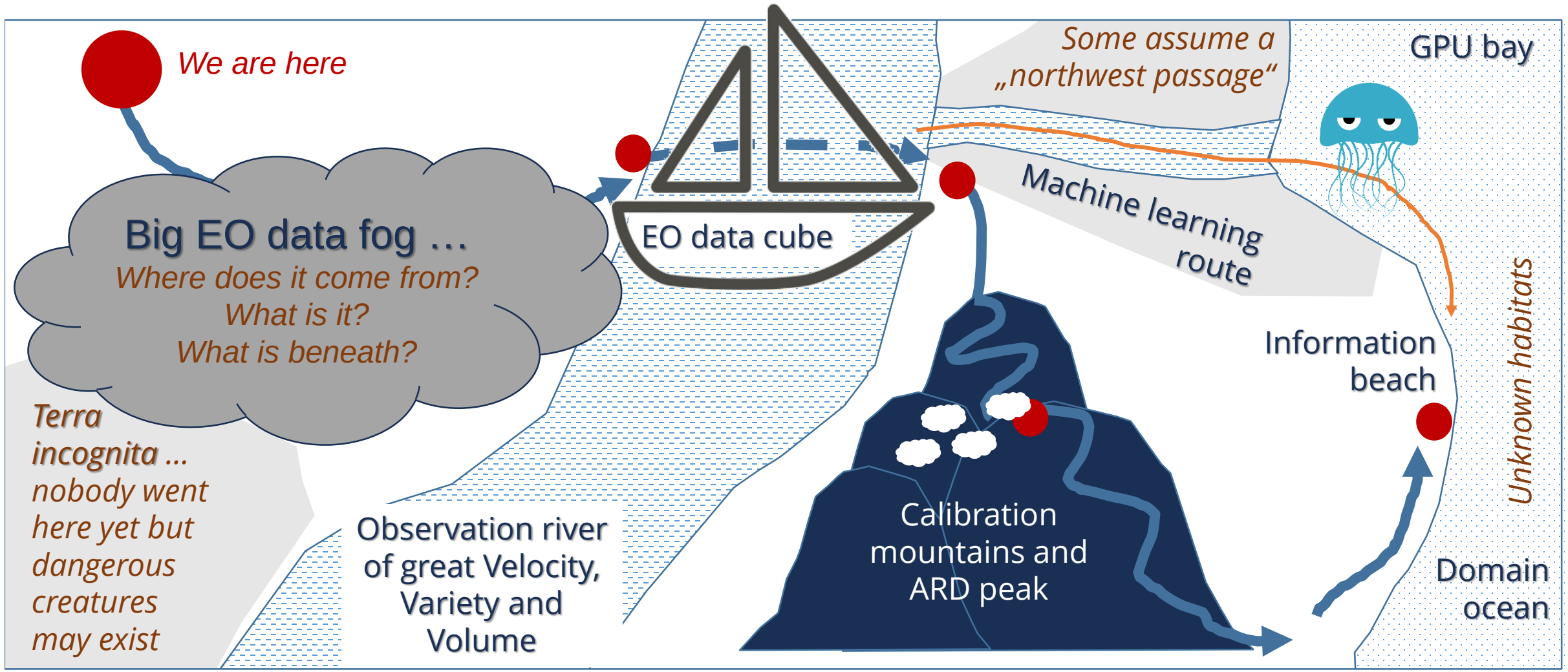
Big EO data analytics - overview

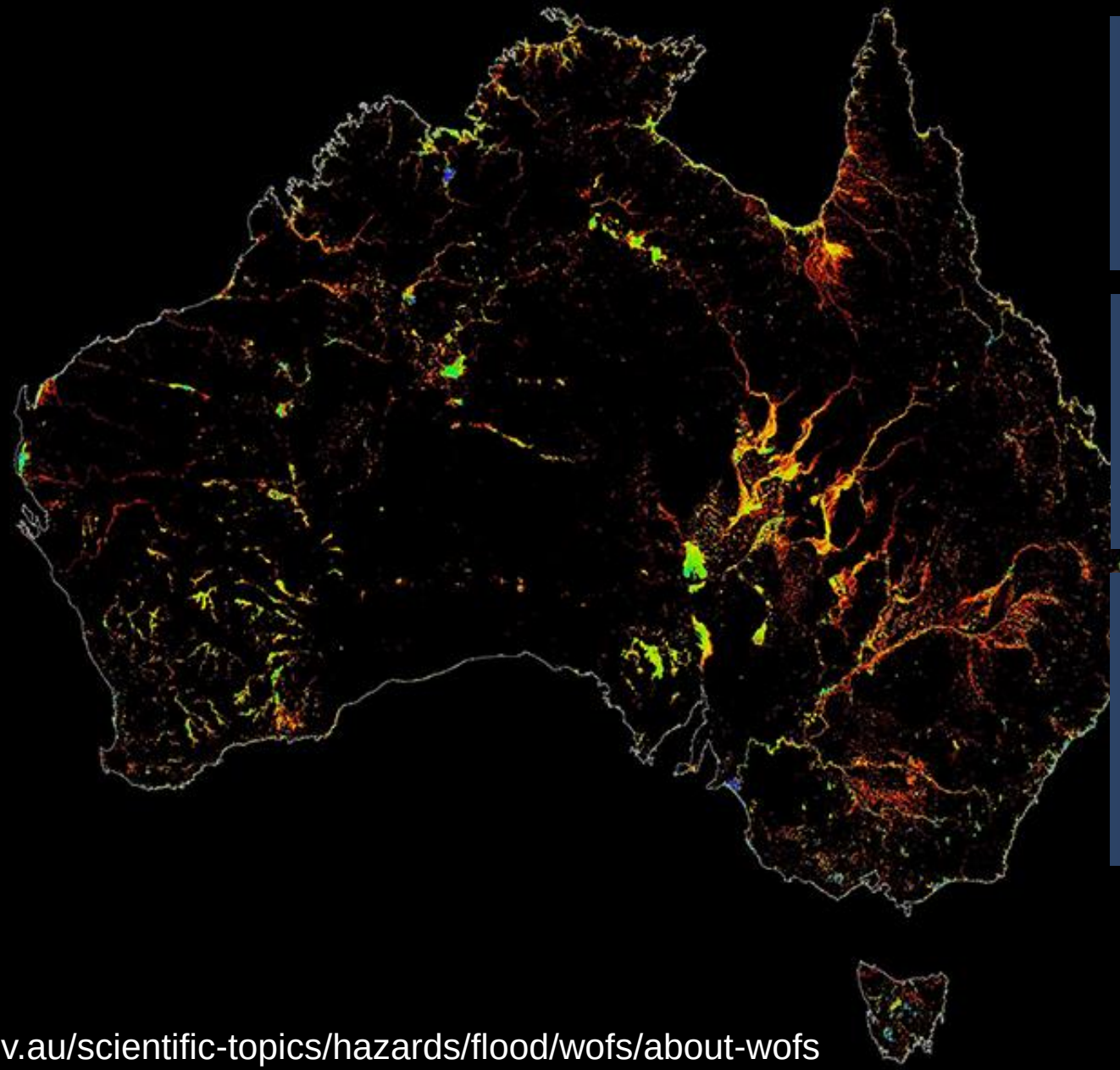
Concepts and technologies for big EO analytics

TAT-7 | 21.06.2019 | Martin Sudmanns, Dirk Tiede, & contrib. Z_GIS colleagues

Interfaculty Department of Geoinformatics – Z_GIS, University of Salzburg

Our „route“ through the Big EO data landscape



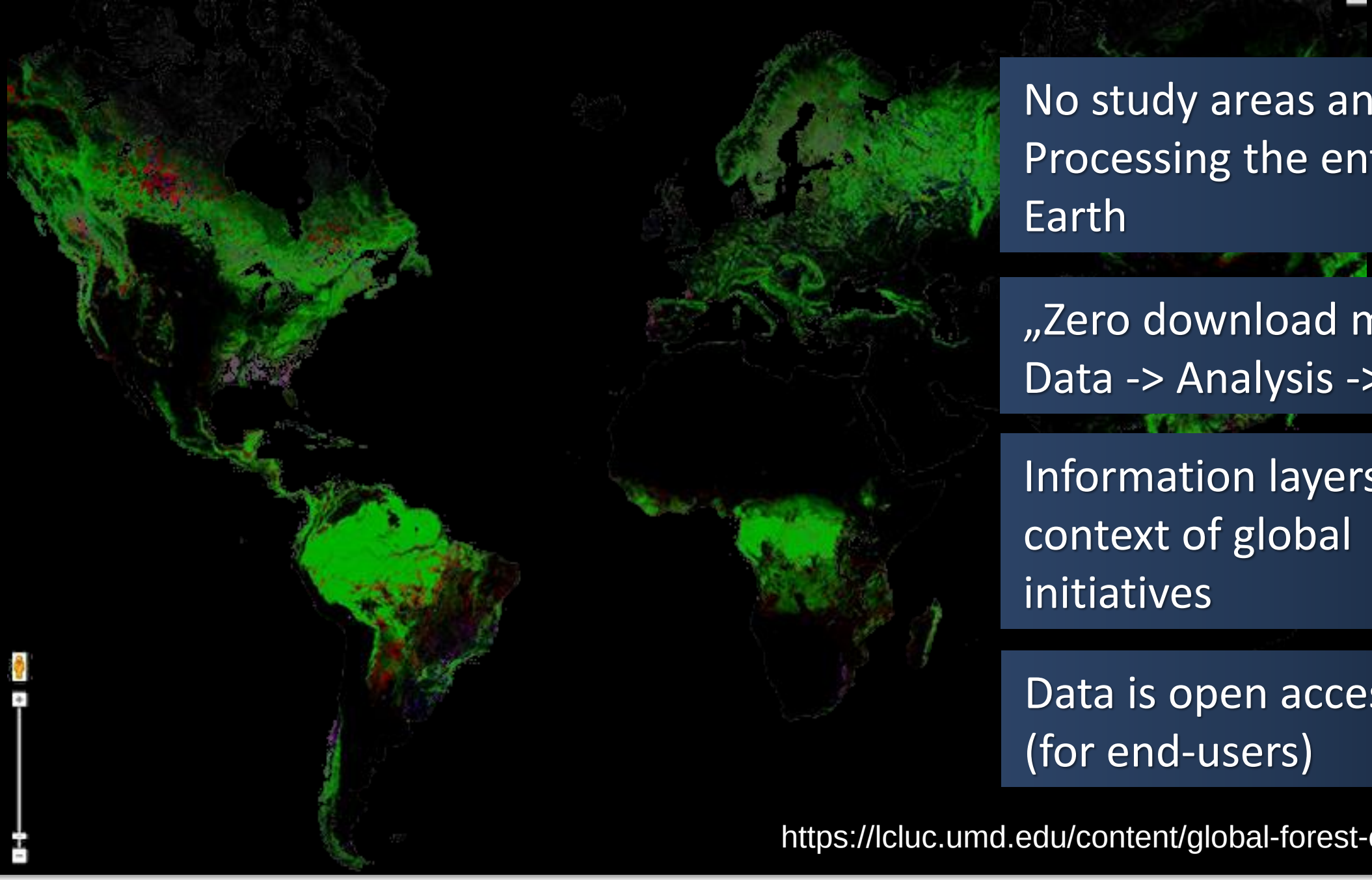


http://www.ga.gov.au/scientific-topics/hazards/flood/wofs/about-wofs

No images any more: Treating pixels as observation in space/time

Executing the algorithm within a cloud over the Internet

No snapshot: “This is a nice map, but how has it changed? ” Mike Wulder: https://www.youtube.com/watch?v=GzIOzTGB-Vs



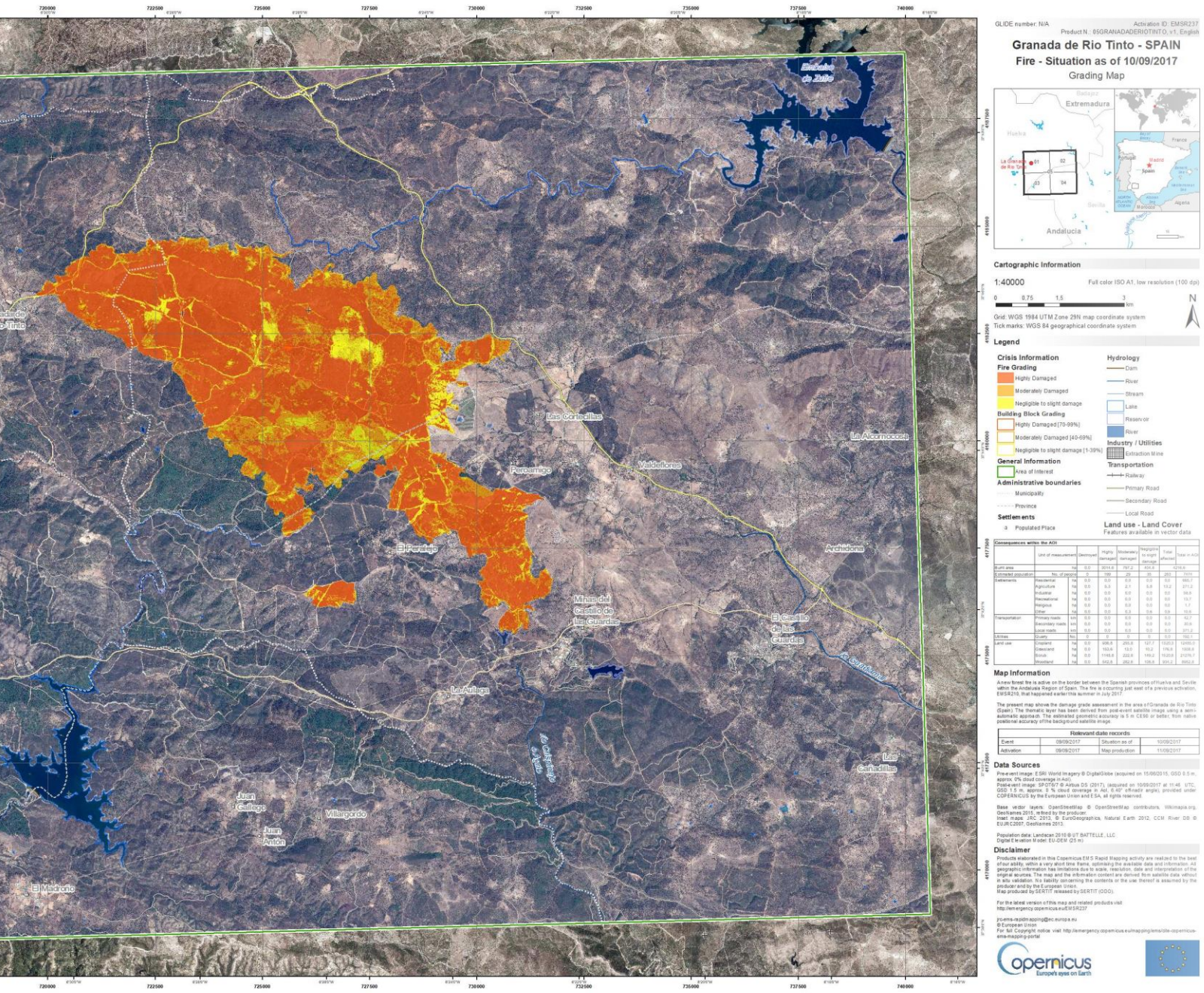
No study areas any more: Processing the entire Earth

„Zero download model“: Data -> Analysis -> Map

Information layers in the context of global initiatives

Data is open access & free (for end-users)

https://lcluc.umd.edu/content/global-forest-change



Information in real-time

General-purpose collection of data

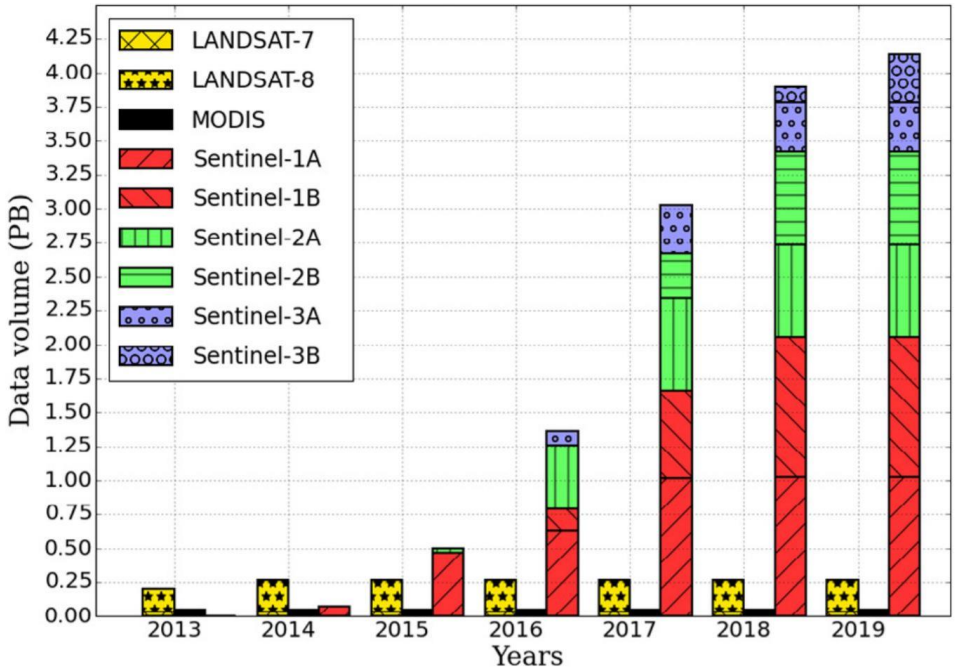
Multi-sensor approach: Virtual constellations

Continous stream of data and update of information

https://emergency.copernicus.eu/mapping/system/files/components/EMSR237_05GRANADADERIOTINTO_02GRADING_MAP_v1_100dpi.jpg

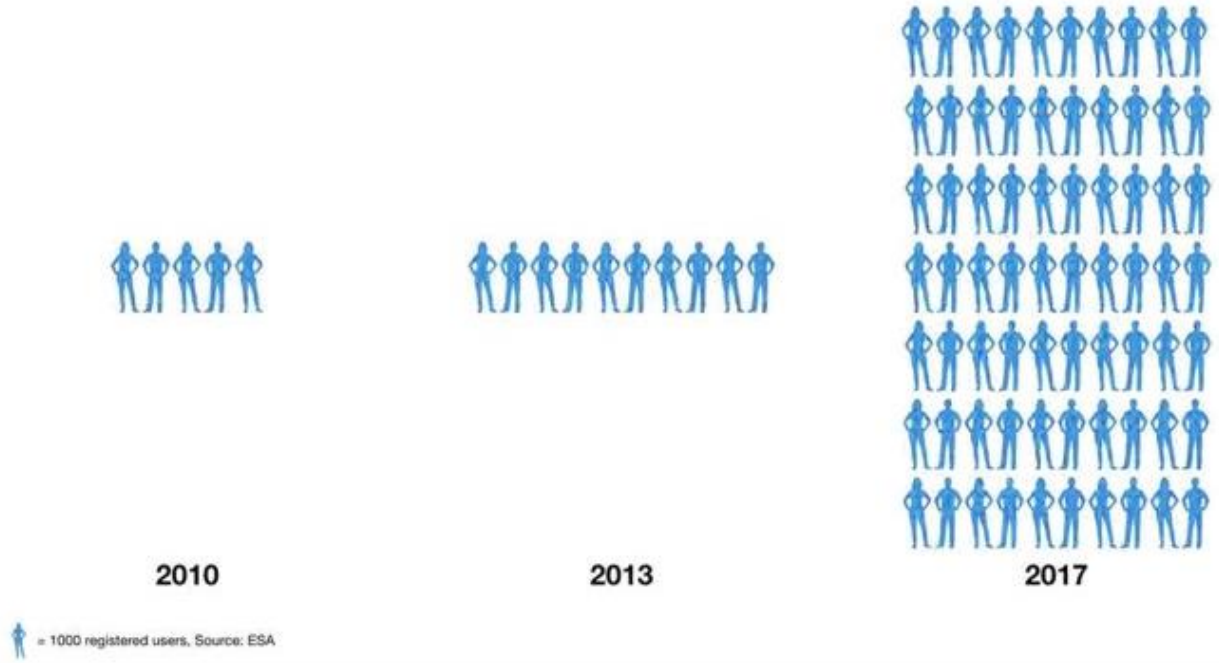
What is big Earth data? – The data perspective

- Unprecedented **volume, variety and velocity** of EO data
- **New user** types and applications



Soille, P., Burger, A., Marchi, D. de, Kempeneers, P., Rodriguez, D., Syrris, V., & Vasilev, V. (2018). A versatile data-intensive computing platform for information retrieval from big geospatial data. Future Generation Computer Systems, 81, 30–40. https://doi.org/10.1016/j.future.2017.11.007

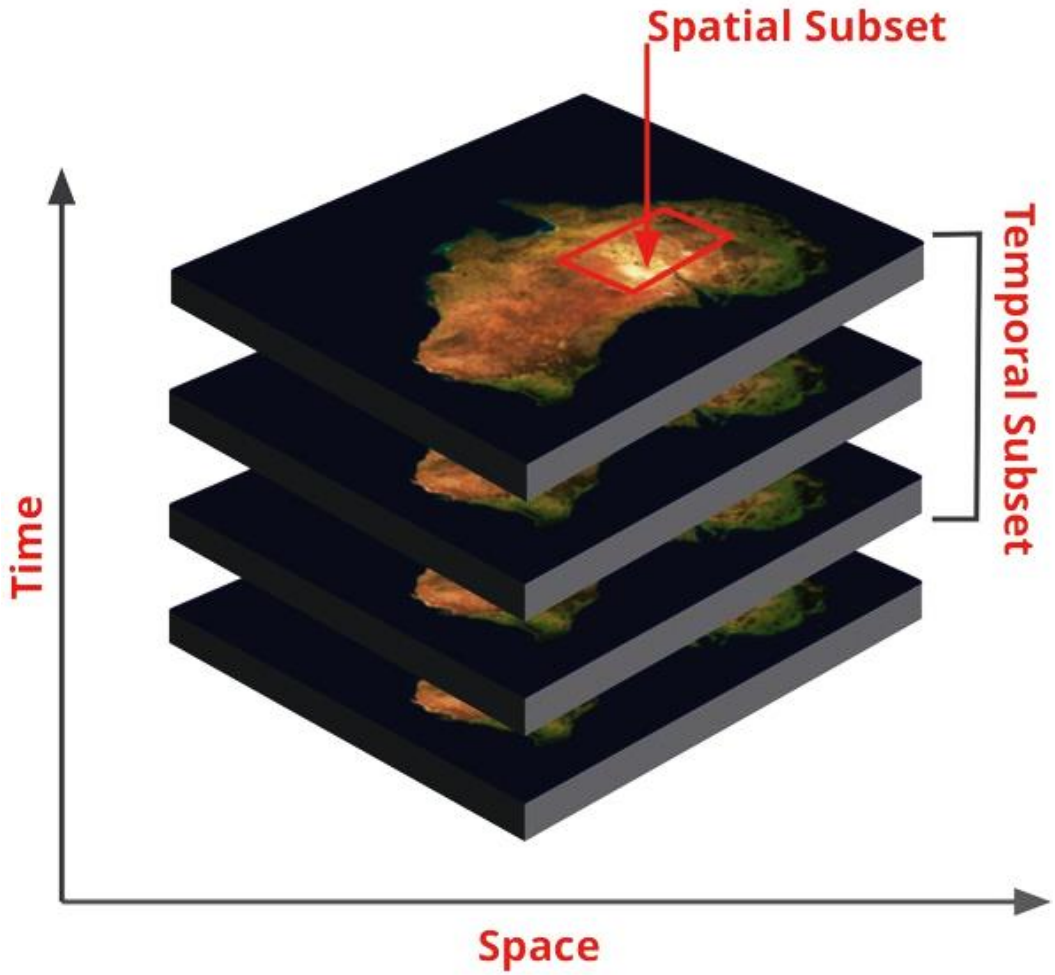
"Bring users to the data, not the data to the users"



Increasing use of EO data, documented by number of registered users at the ESA portals

What is big Earth Observation data?

- Observations instead of images
- Data acquisition not for a single purpose, but **constant stream**
- Image archive date back some **decades** (e.g., MODIS, Landsat)
- Availability of **computing power**
- Increased **acquisition frequency**
- Decreased time between acquisition and data provision
- Data are **open access**
- Volume, velocity, variety



Sudmanns, M., Lang, S. and Tiede, D., Big Earth Data: From Data to Information. GI_Forum 2018,, 1, pp.184-193.

4.1

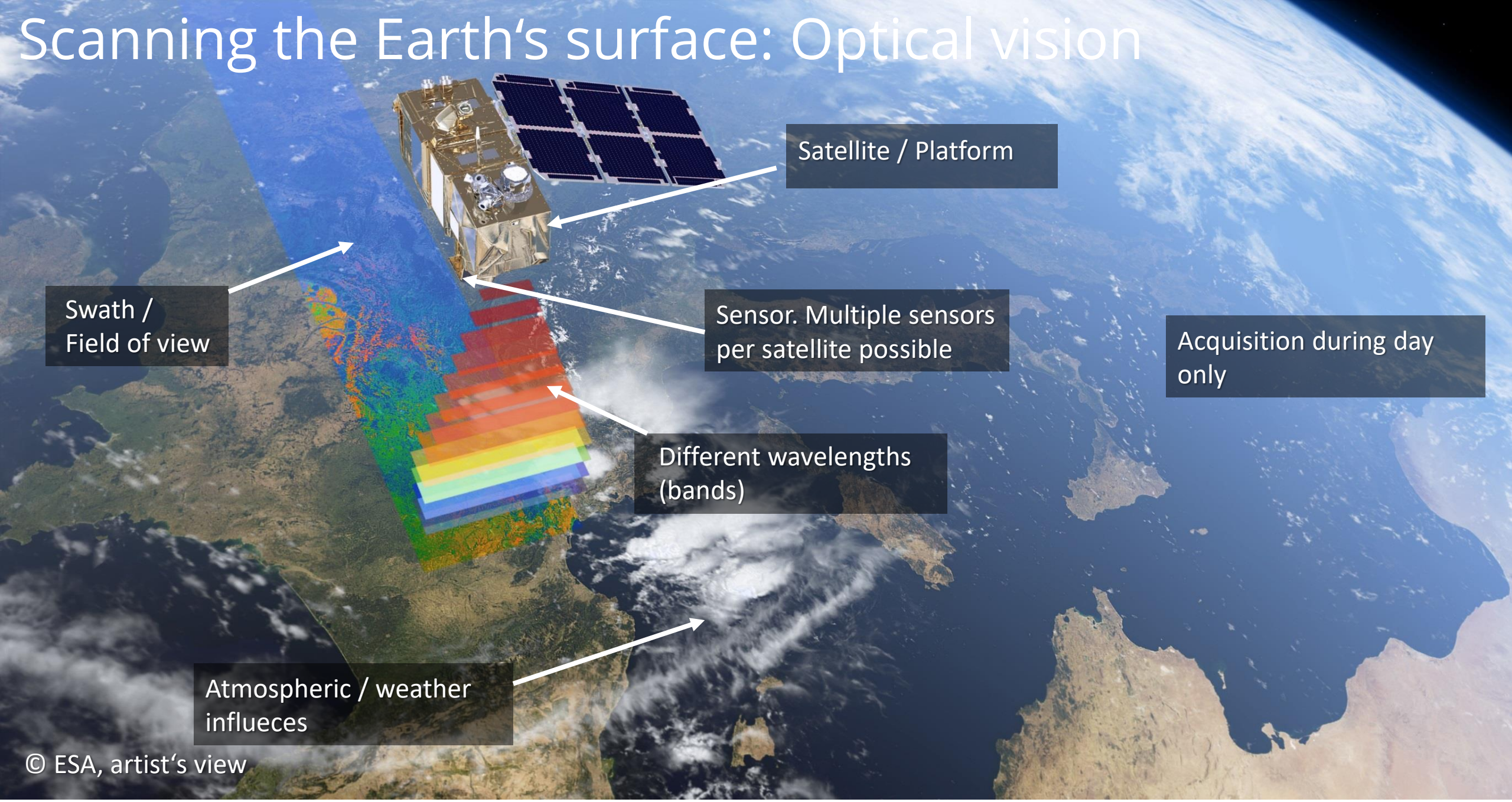
The “Mechanics” of EO

Concepts and technologies for big EO analytics

TAT-7 | 21.06.2019 | Martin Sudmanns, Dirk Tiede, & contrib. Z_GIS colleagues

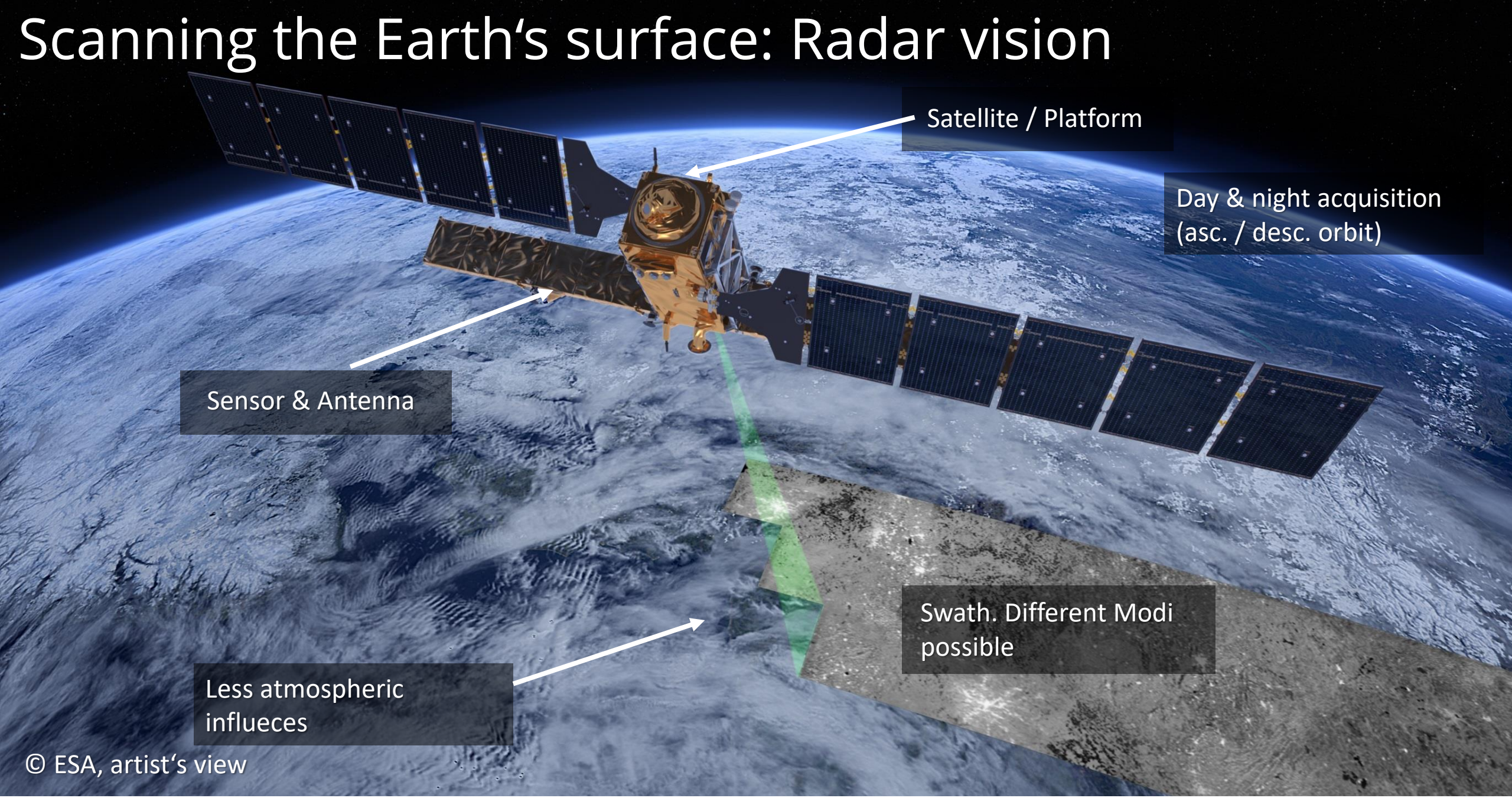
Interfaculty Department of Geoinformatics – Z_GIS, University of Salzburg

Scanning the Earth‘s surface: Optical vision



© ESA, artist‘s view

Scanning the Earth‘s surface: Radar vision



From space to Earth: Where big EO data "happens" …

Space segment:

- Many problems, but different topic …

Ground segment (1):

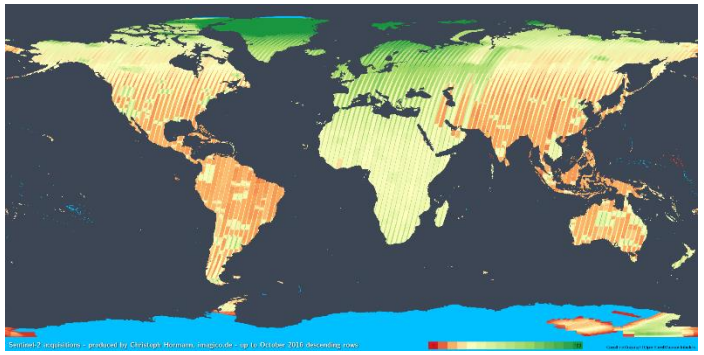
- Lots of technical decisions (input & output, reprocessing), calibration
- Defining processing baseline

Ground segment (2) / data provider:

- Providing data, e.g., in download portals
- Also: using exploitation platforms and online processing

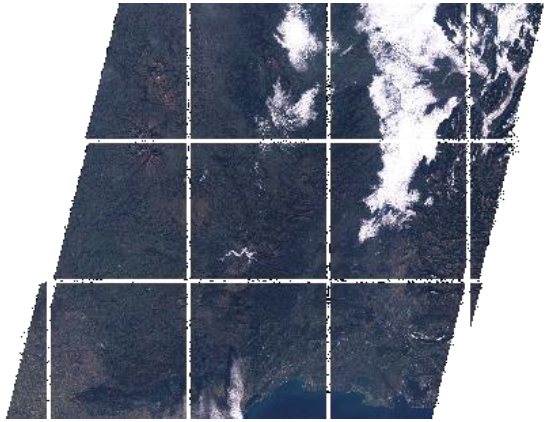
End user:

- Creating value / information from data

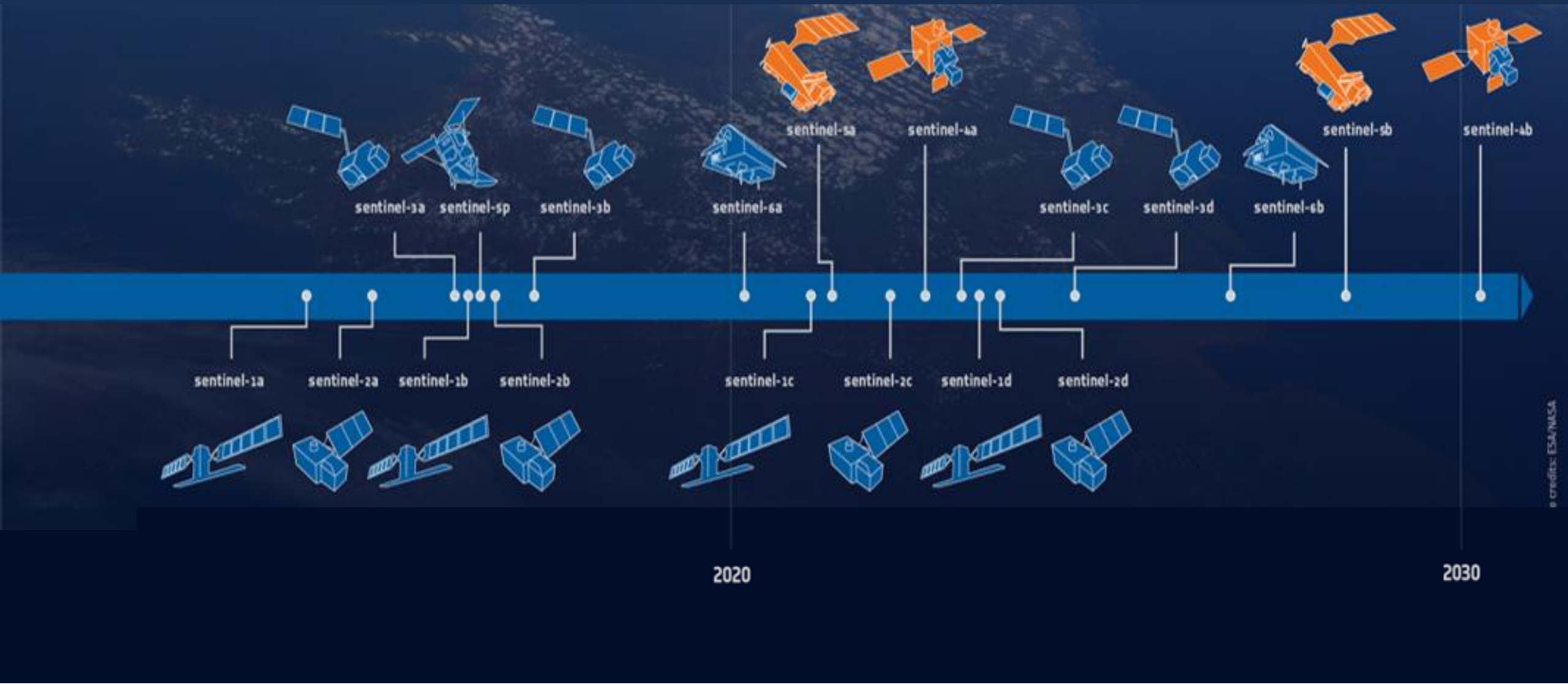


Raw data stream

Image file



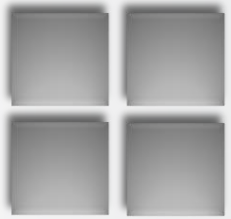
Temporal scale of EO missions



Data deluge: Sentinel-2 example

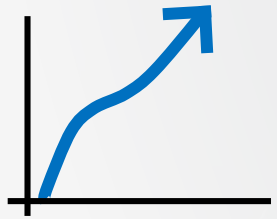
> 6 mil.

Sentinel-2 Scenes in the archive



3.4 TB

new data volume every day

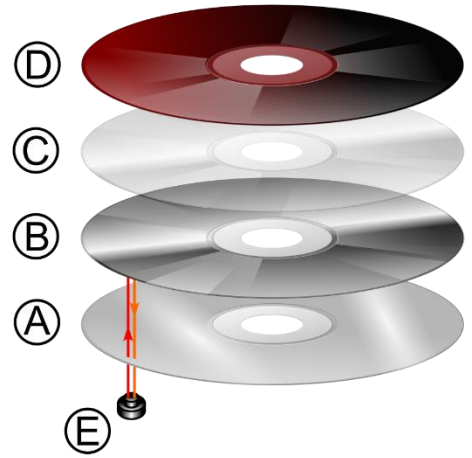


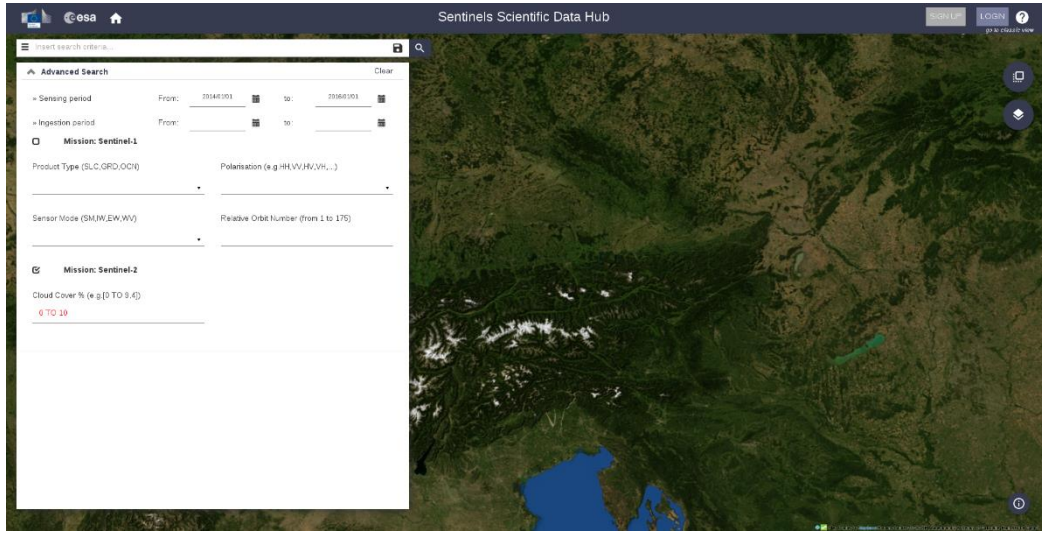
70 %

of the Earth is cloud covered

History of data access

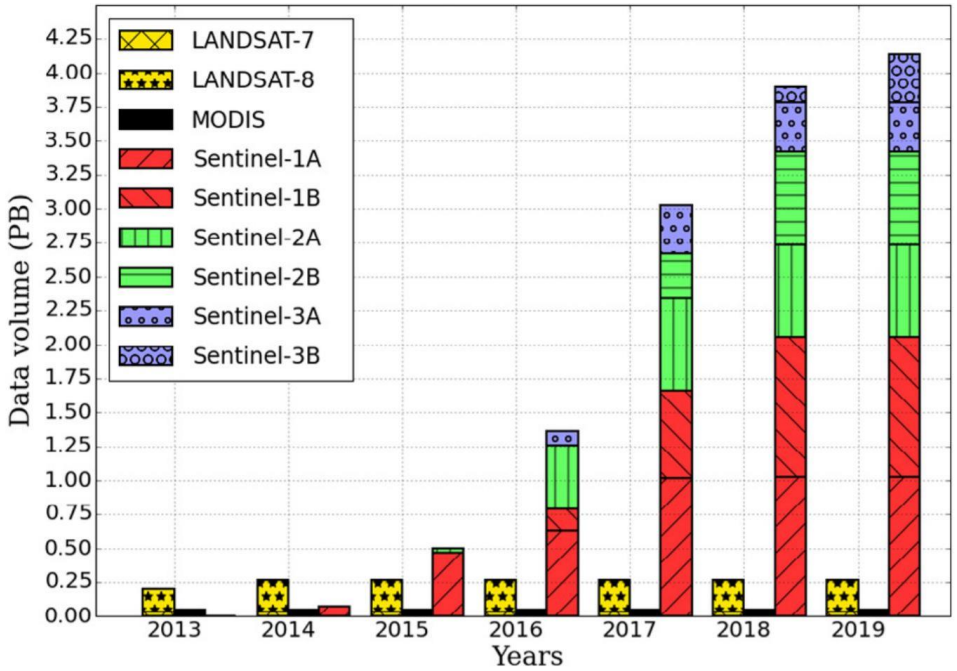
1. Tape by mail
2. CD by mail
3. DVD by mail
4. Download over Internet





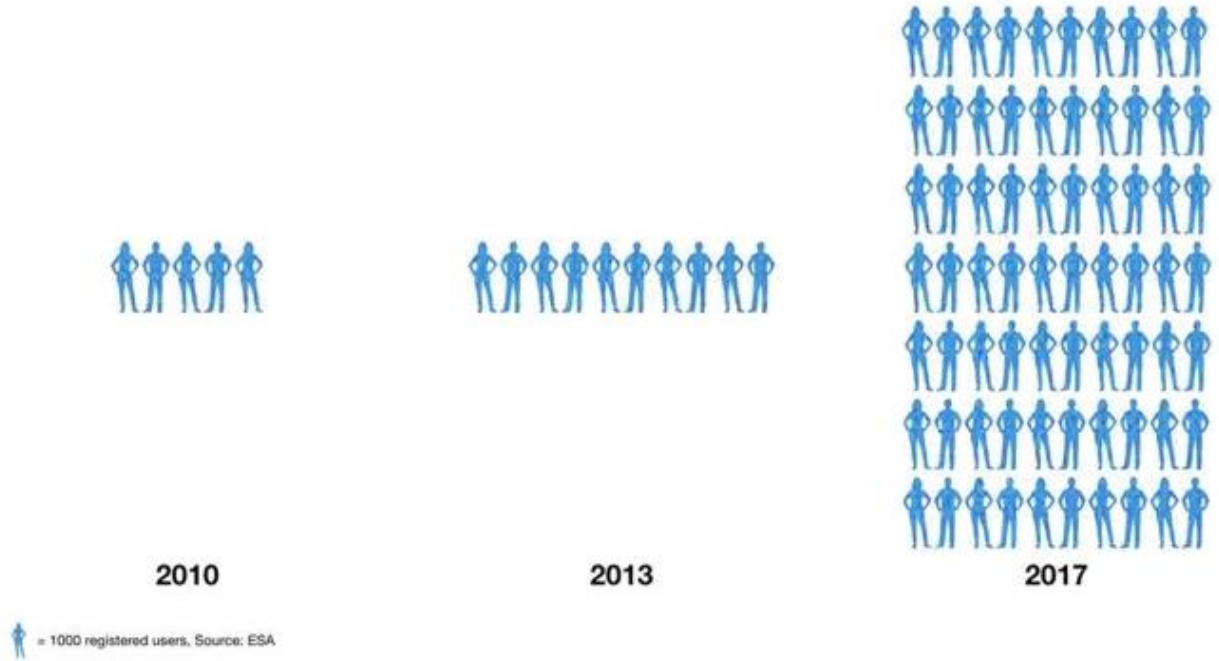
What is big Earth data?

- Unprecedented **volume, variety and velocity** of EO data
- **New user** types and applications



Soille, P., Burger, A., Marchi, D. de, Kempeneers, P., Rodriguez, D., Syrris, V., & Vasilev, V. (2018). A versatile data-intensive computing platform for information retrieval from big geospatial data. Future Generation Computer Systems, 81, 30–40. https://doi.org/10.1016/j.future.2017.11.007

"Bring users to the data, not the data to the users"

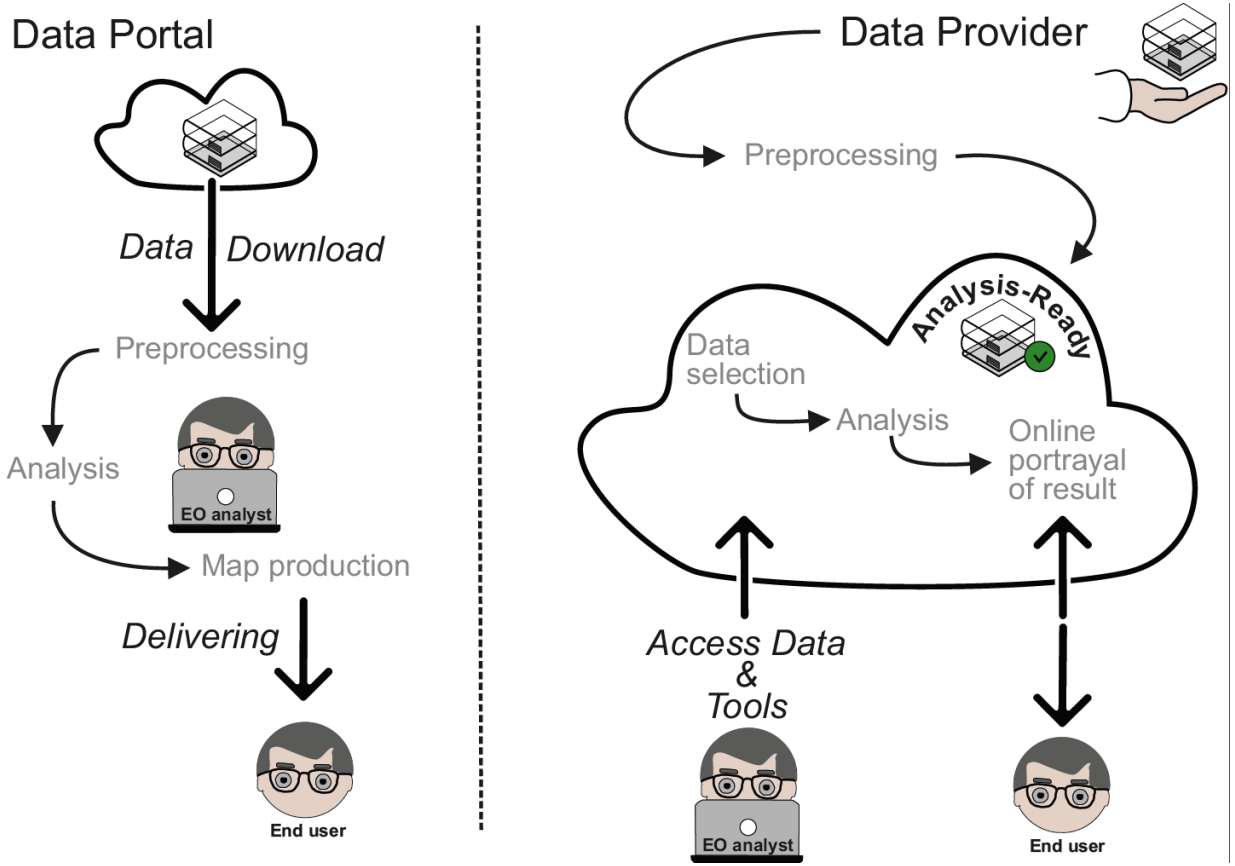


Increasing use of EO data, documented by number of registered users at the ESA portals

History of data access

1. Tape by mail
2. CD by mail
3. DVD by mail
4. Download over Internet
5. **Data do not move: "Bring user to the data, not data to the user"**

„Zero download model"

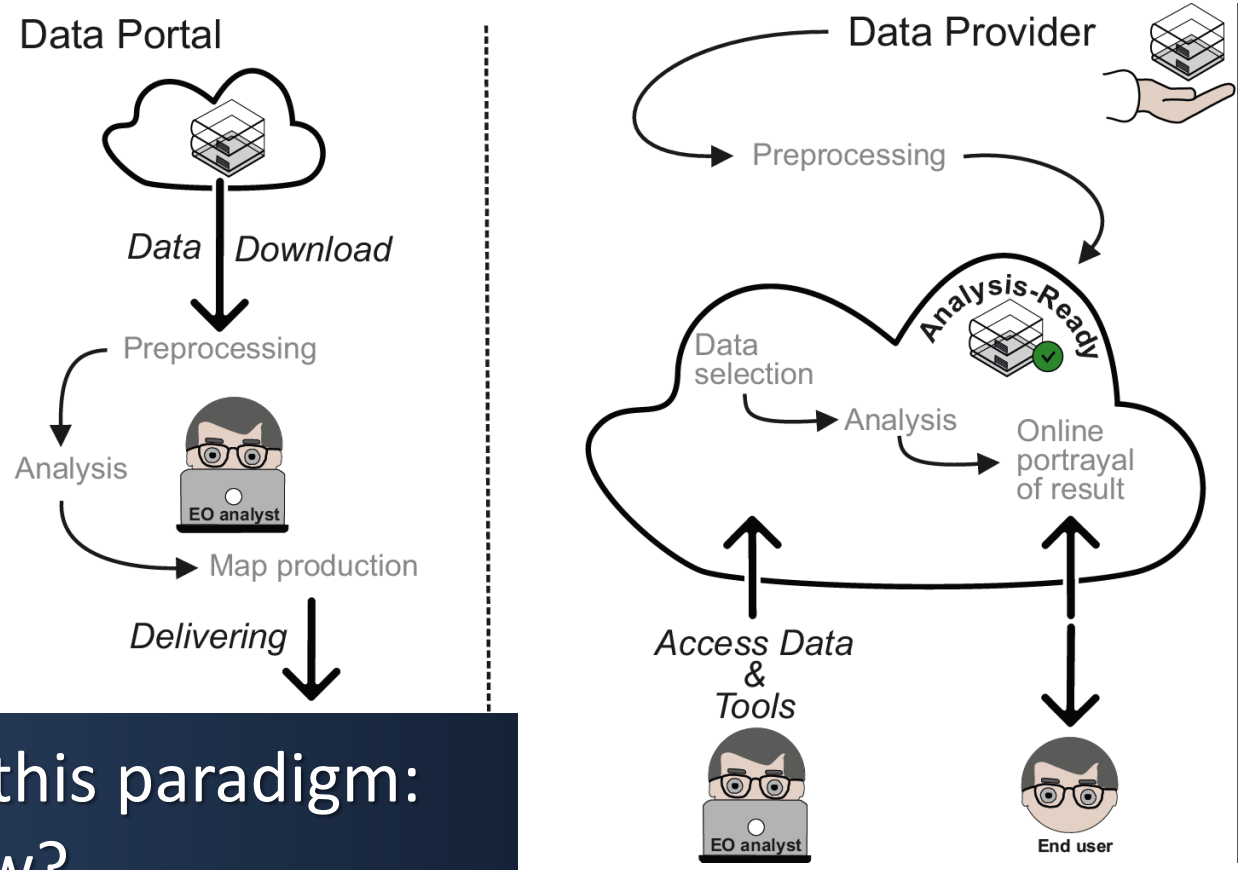


Sudmanns, M., Tiede, D., Lang, S., Bergstedt, H., Trost, G., Augustin, H., Baraldi, A. and Blaschke, T., 2019. Big Earth data: disruptive changes in Earth observation data management and analysis?. International Journal of Digital Earth, pp.1-19.

History of data access

1. Tape by mail
2. CD by mail
3. DVD by mail
4. Download over Internet
5. **Data do not move: "Bring user to the data, not data to the user"**

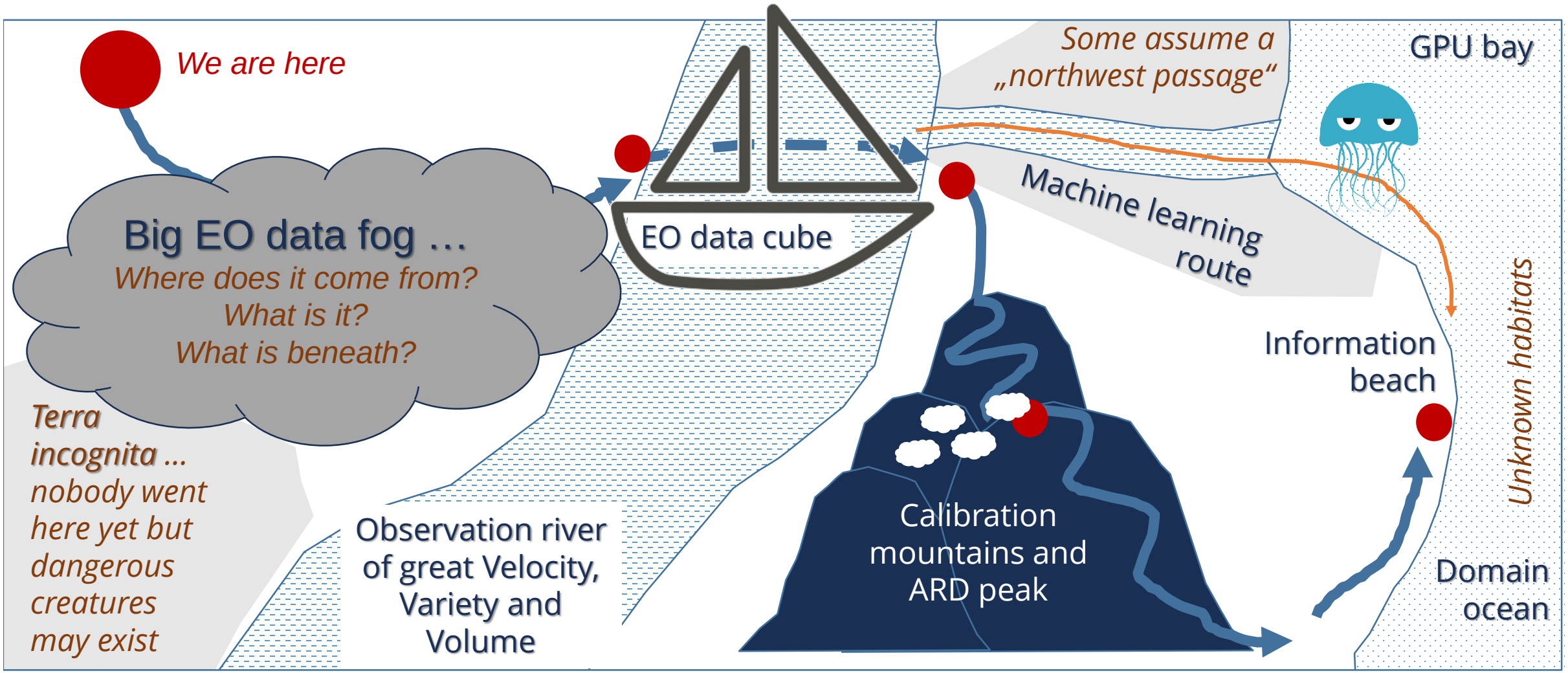
„Zero download model"



Implementing this paradigm: Question is how?

, Bergstedt, H., Trost, G., Augustin, H., Baraldi, A. ta: disruptive changes in Earth observation data management and analysis?. International Journal of Digital Earth, pp.1-19.

Our „route“ through the Big EO data landscape



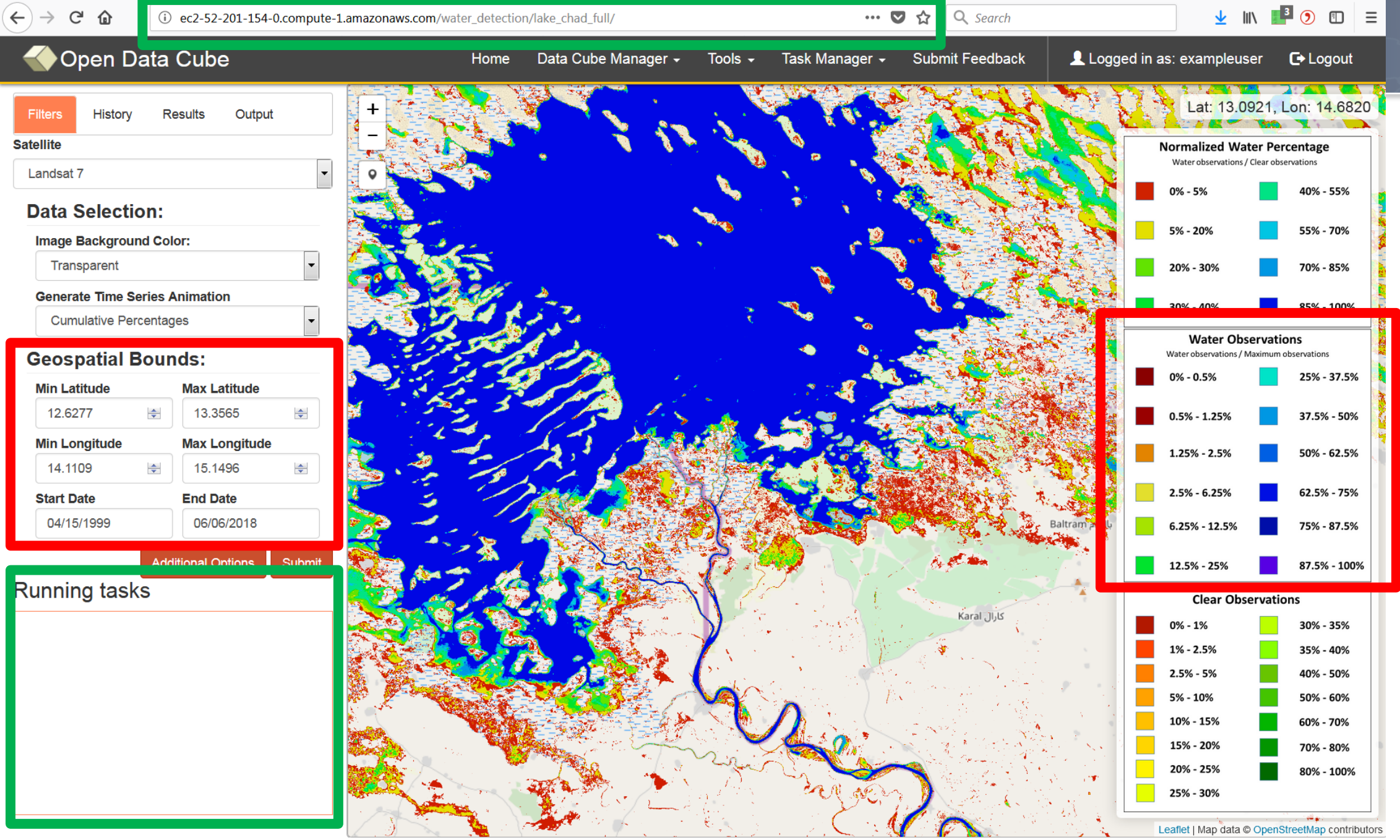
4.2

Earth observation data cubes

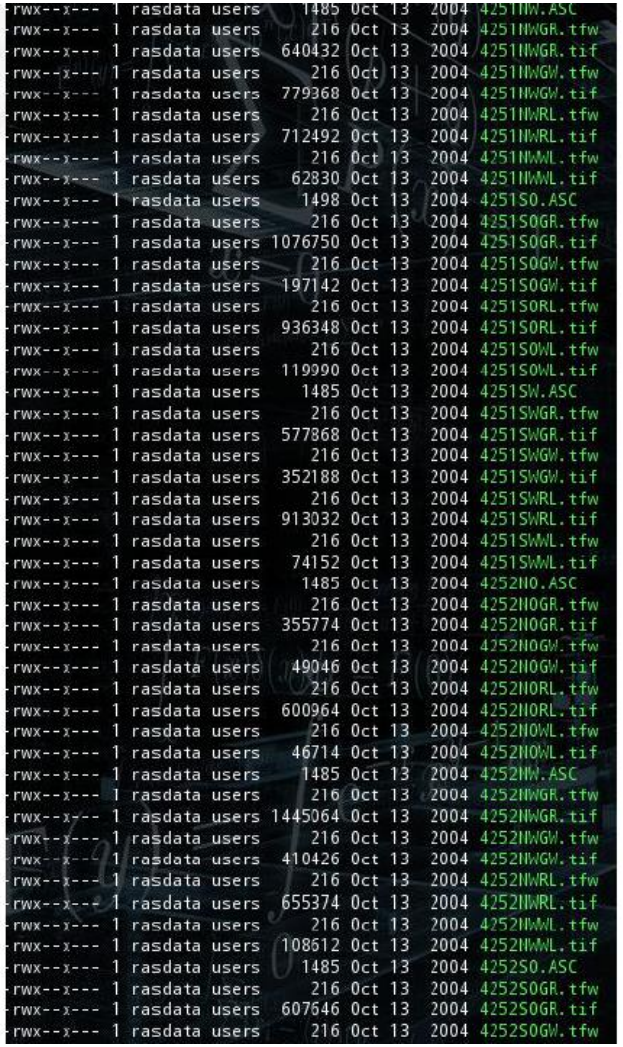
Concepts and technologies for big EO analytics

TAT-7 | 21.06.2019 | Martin Sudmanns, Dirk Tiede, & contrib. Z_GIS colleagues

Interfaculty Department of Geoinformatics – Z_GIS, University of Salzburg


ec2-52-201-154-0.compute-1.amazonaws.com/water_detection/lake_chad_full/
Search
Open Data Cube
Home
Data Cube Manager
Tools
Task Manager
Submit Feedback
Logged in as: exampleuser
Logout
Filters
History
Results
Output
Satellite
Landsat 7
Data Selection:
Image Background Color:
Transparent
Generate Time Series Animation
Cumulative Percentages
Geospatial Bounds:
Min Latitude
12.6277
Max Latitude
13.3565
Min Longitude
14.1109
Max Longitude
15.1496
Start Date
04/15/1999
End Date
06/06/2018
Additional Options
Submit
Running tasks
Lat: 13.0921, Lon: 14.6820
Normalized Water Percentage
Water observations / Clear observations
0% - 5%
5% - 20%
20% - 30%
40% - 55%
55% - 70%
70% - 85%
Water Observations
Water observations / Maximum observations
0% - 0.5%
0.5% - 1.25%
1.25% - 2.5%
2.5% - 6.25%
6.25% - 12.5%
12.5% - 25%
25% - 37.5%
37.5% - 50%
50% - 62.5%
62.5% - 75%
75% - 87.5%
87.5% - 100%
Clear Observations
0% - 1%
1% - 2.5%
2.5% - 5%
5% - 10%
10% - 15%
15% - 20%
20% - 25%
25% - 30%
30% - 35%
35% - 40%
40% - 50%
50% - 60%
60% - 70%
70% - 80%
80% - 100%
Baltram
Karal
Leaflet | Map data © OpenStreetMap contributors

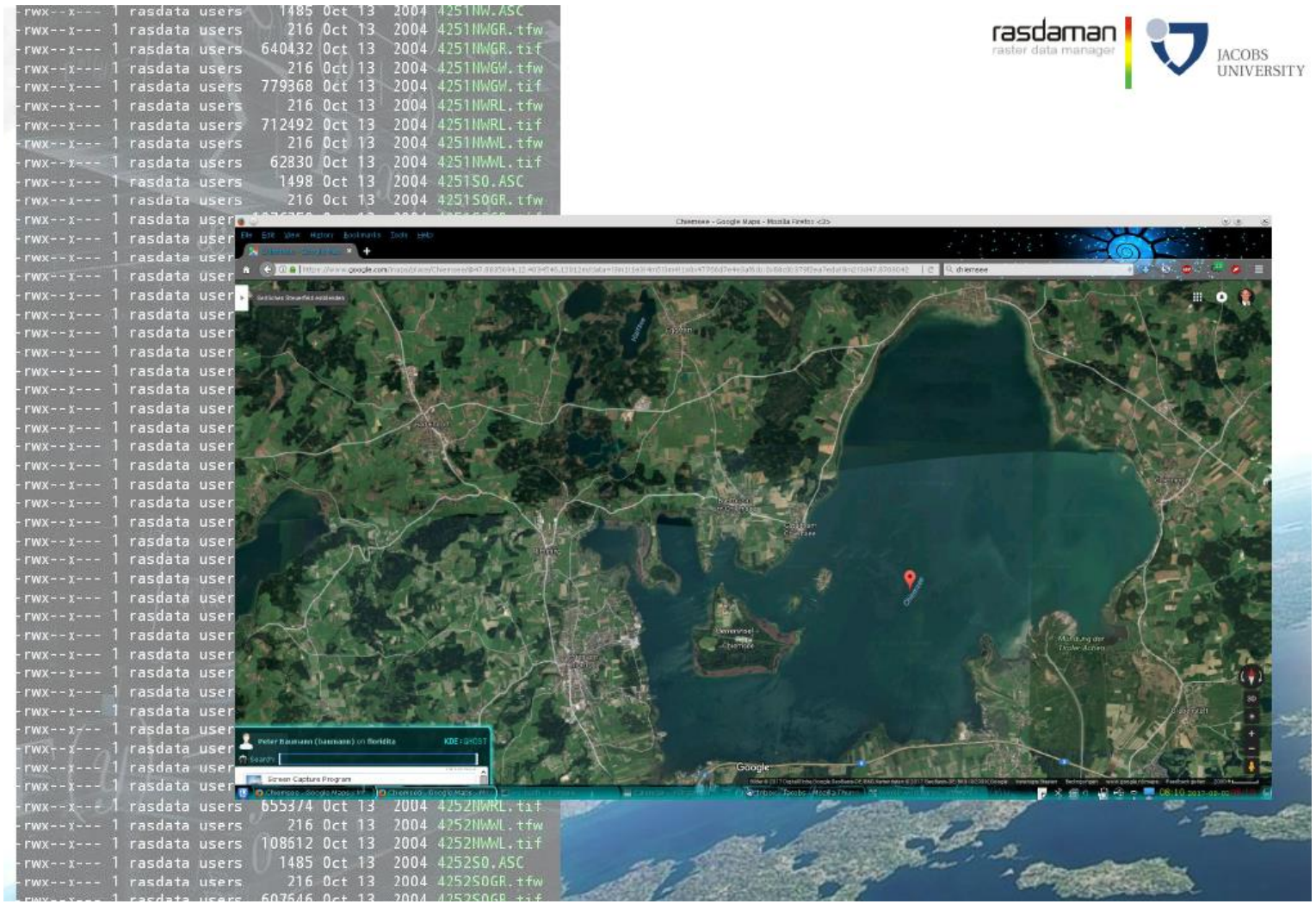
Geospatial Data Cubes



rasdaman raster data manager | JACOBS UNIVERSITY

Based on material by Peter Baumann

Geospatial Data Cubes



rasdaman raster data manager | JACOBS UNIVERSITY

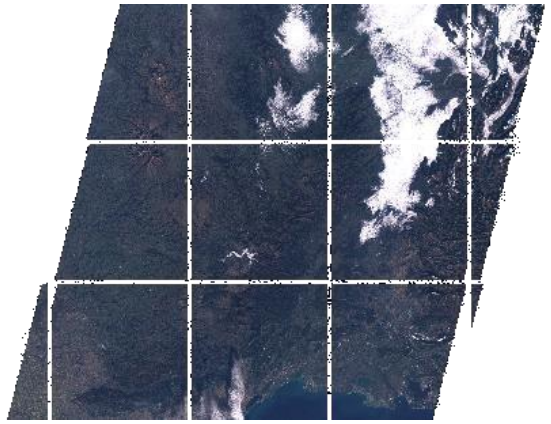
How do you store your pictures from your last vacation?

How does Google Earth store the images?

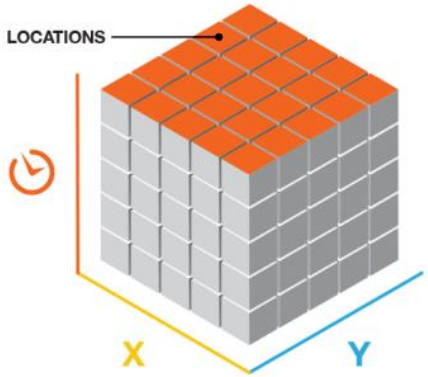
Based on material by Peter Baumann

Thinking in observations instead of images

Aquisitions, raw data stream



Distribution to users as images



Centralised storage of observations in space-time

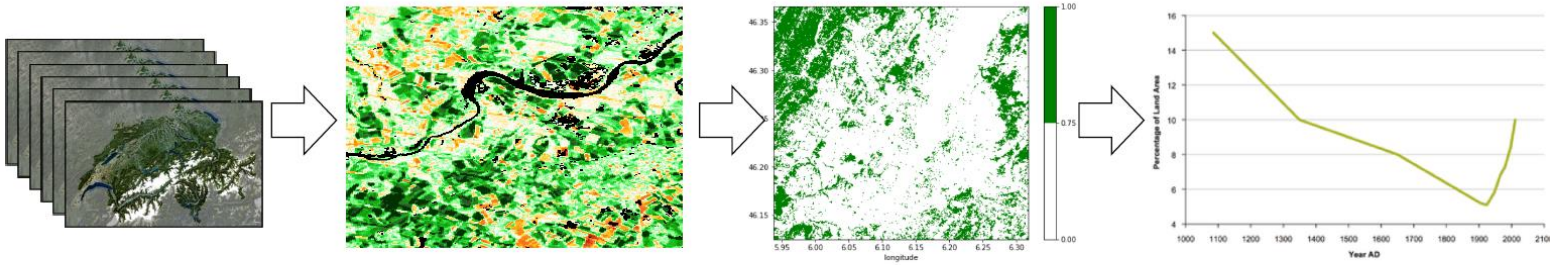
http://desktop.arcgis.com/en/arcmap/10.3/tools/space-time-pattern-mining-toolbox/learnmorecreatecube.htm

Image-based approach, temporal snapshots

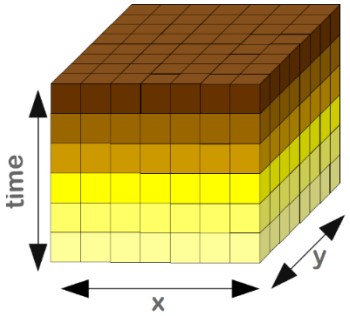
Data cubes: No images; at least one non-spatial axis

Geospatial Data Cubes

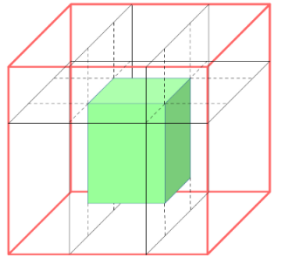
- Store data **query-optimised not acquisition-oriented**



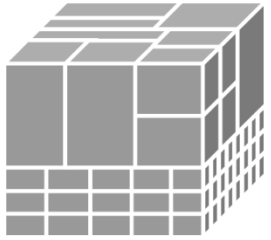
From: http://www.swissdatacube.org/



1.) Voxels arranged as dense temporal stack



2.) Data access using declarative query language



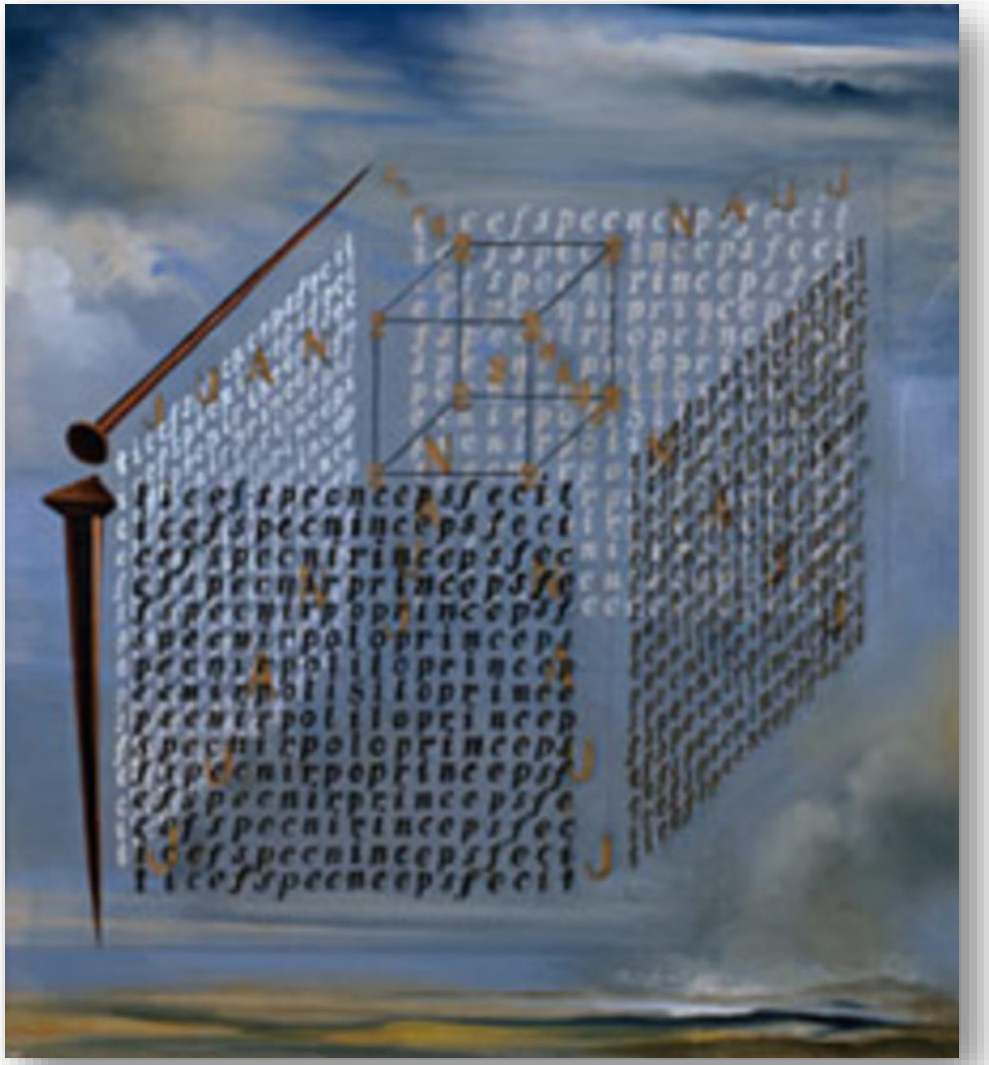
3.) Regular or irregular partitioning (tiling)

- Different access methods (API, query language)
- Data cubes as infrastructure
- Provide a **logical view** on the data

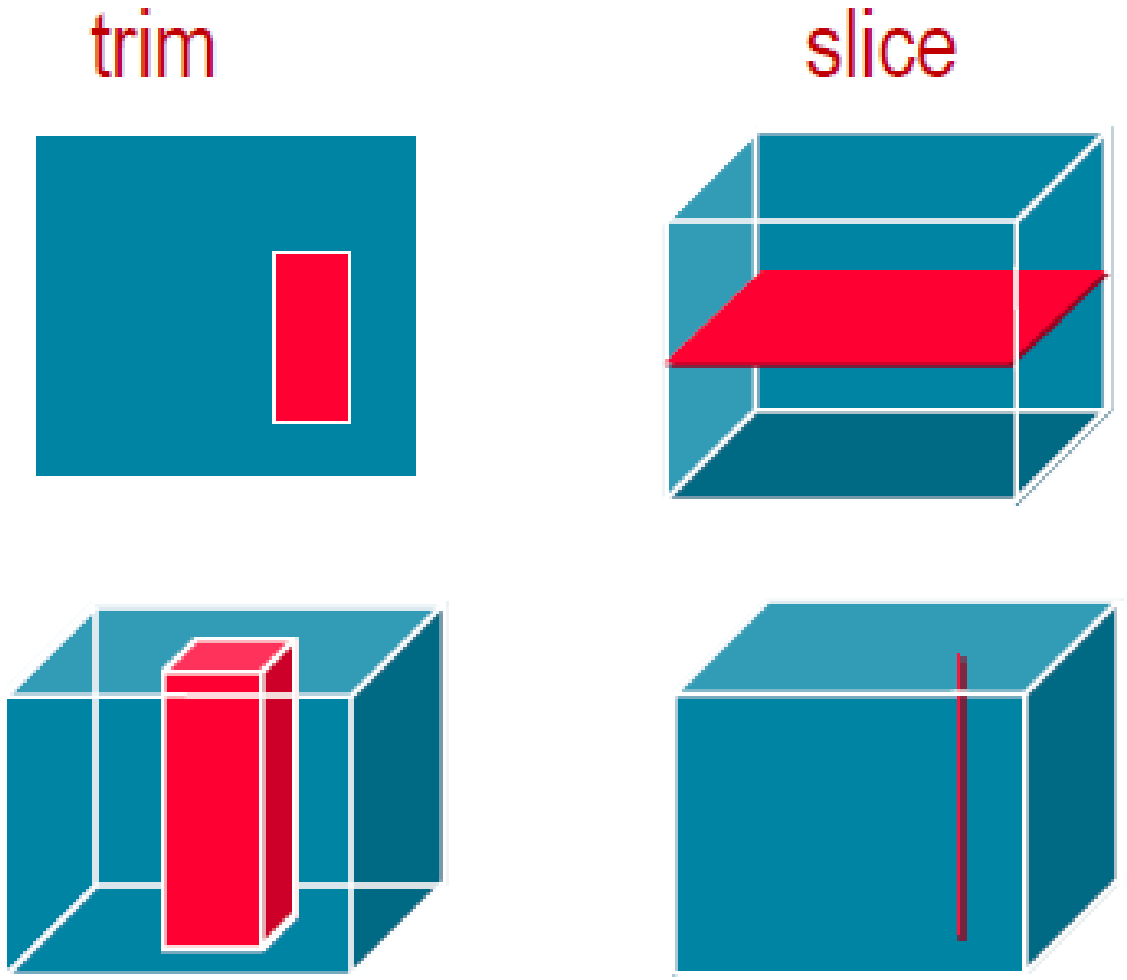
1. Index external files
2. data as multi-dimensional array

Tiede, Dirk; Baraldi, Andrea; Sudmanns, Martin; Belgiu, Mariana; Lang, Stefan (2017): Architecture and prototypical implementation of a semantic querying system for big Earth observation image bases. In European journal of remote sensing 50 (1), pp. 452–463. DOI: 10.1080/22797254.2017.1357432.

Geospatial Data Cubes



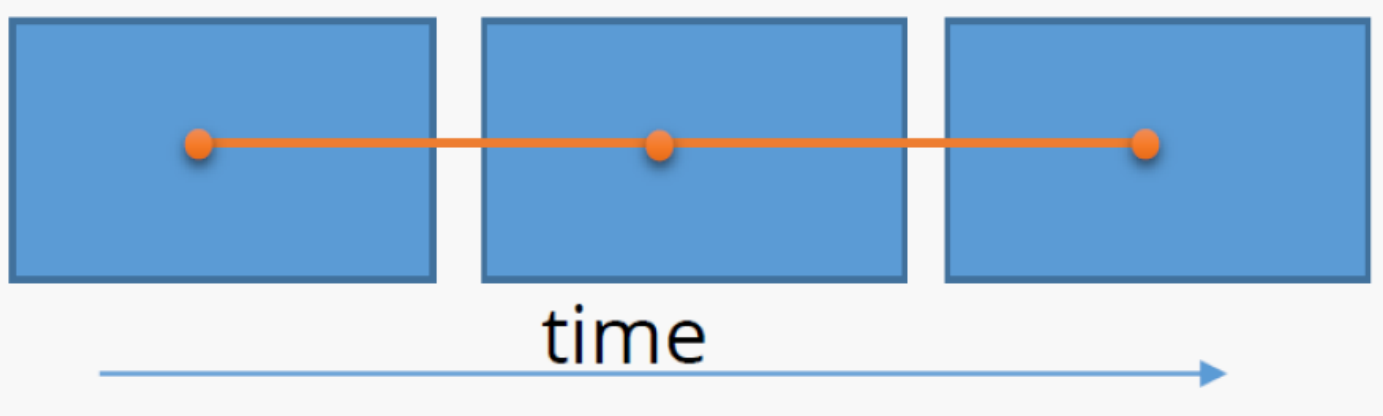
Salvador Dali (1960): "A Propos of the 'Treatise on Cubic Form' by Juan de Herrera"



Operations in a 3D array to trim or slice a data cube

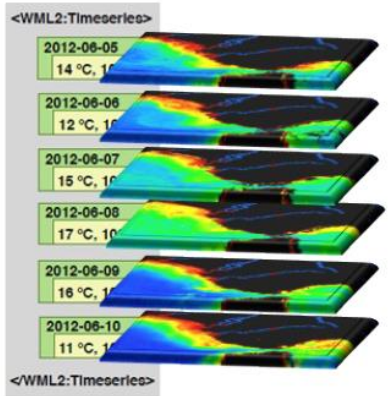
Space first vs. time first

- With time series analysis, the reading “direction” is opposite to the storing “direction” with images
- Using data more often (n) than acquisition (1)
- Example: 100 images with 120 Mio. pixel / image
  → reading 1.2 Bil. pixel, even for a 100 pixel time series

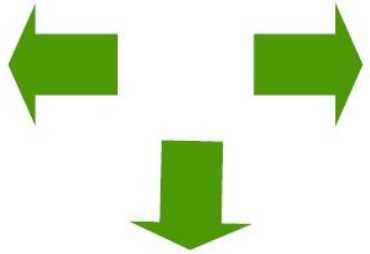


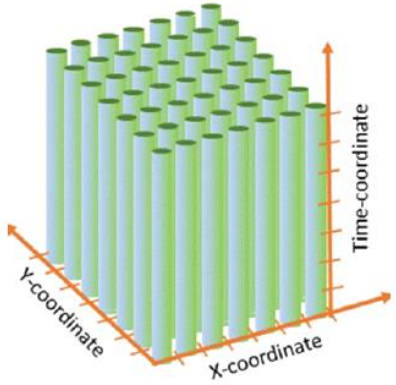
Space first vs. time first

- There is no optimal solution*, only trade-offs

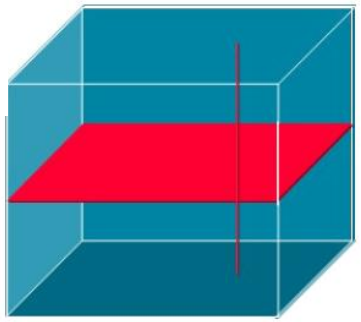


Stacked images: WaterML, GeoServer, most WebGIS applications





Time series, e.g. NASA data rods



*which serves all applications

Data Cube Manifesto

A definition of a data cube

> *"A datacube is a massive multi-dimensional array, also called "raster data" or "gridded data"; "massive" entails that we talk about sizes significantly beyond the main memory resources of the server hardware. Data values, all of the same data type, sit at grid points as defined by the d axes of the d dimensional datacube. Coordinates along these axes allow addressing data values unambiguously.*
>
> *A d-dimensional grid is characterized by the fact that each inner grid point has exactly two neighbors along each direction; border grid points have just one. Point clouds, e.g., are not grids."*

http://www.earthserver.eu/tech/datacube-manifesto

Open Data Cube Perspective

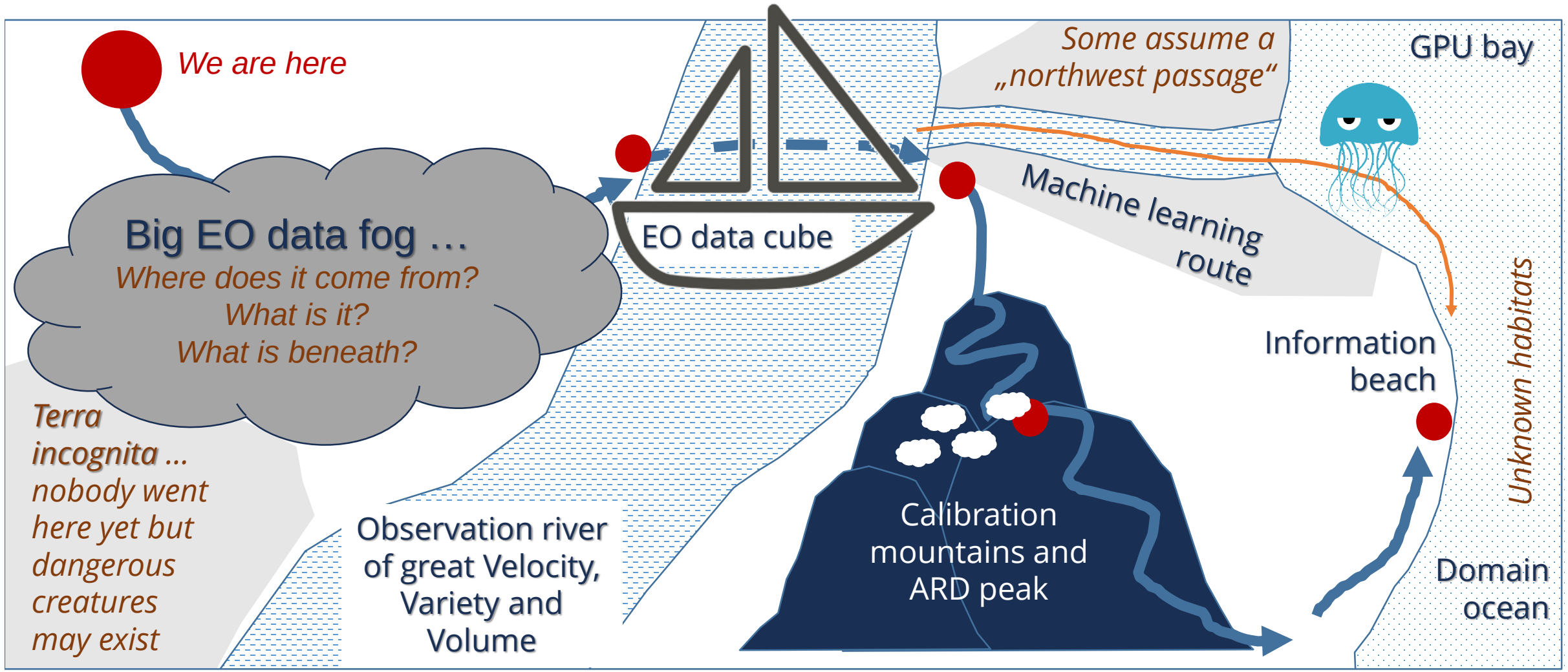
Open Data Cube Perspective

The objective of the ODC is to increase the impact of satellite data by providing an open and freely accessible exploitation tool, and to foster a community to develop, sustain, and grow the breadth and depth of applications.
http://opendatacube.org

Open Data Cube Perspective

Emerging in several countries → Training in exercise 3!

Our „route“ through the Big EO data landscape



4.3

Analysis-ready data (ARD)

Concepts and technologies for big EO analytics

TAT-7 | 21.06.2019 | Martin Sudmanns, Dirk Tiede, & contrib. Z_GIS colleagues

Interfaculty Department of Geoinformatics – Z_GIS, University of Salzburg

Analysis-ready data (ARD)

- Analysts spend too much time in **data preparation**
- The community seeks a solution in providing data that can be **direct input to analysis**
- Obstacles can be
  - EO data are complicated to use
  - Data discovery & data access
  - File format and geometric transformation
  - Incompatible processing levels and pre-processing
  - …

https://medium.com/planet-stories/analysis-ready-data-defined-5694f6f48815

Analysis-ready data (ARD)

Existing definitions (selection)

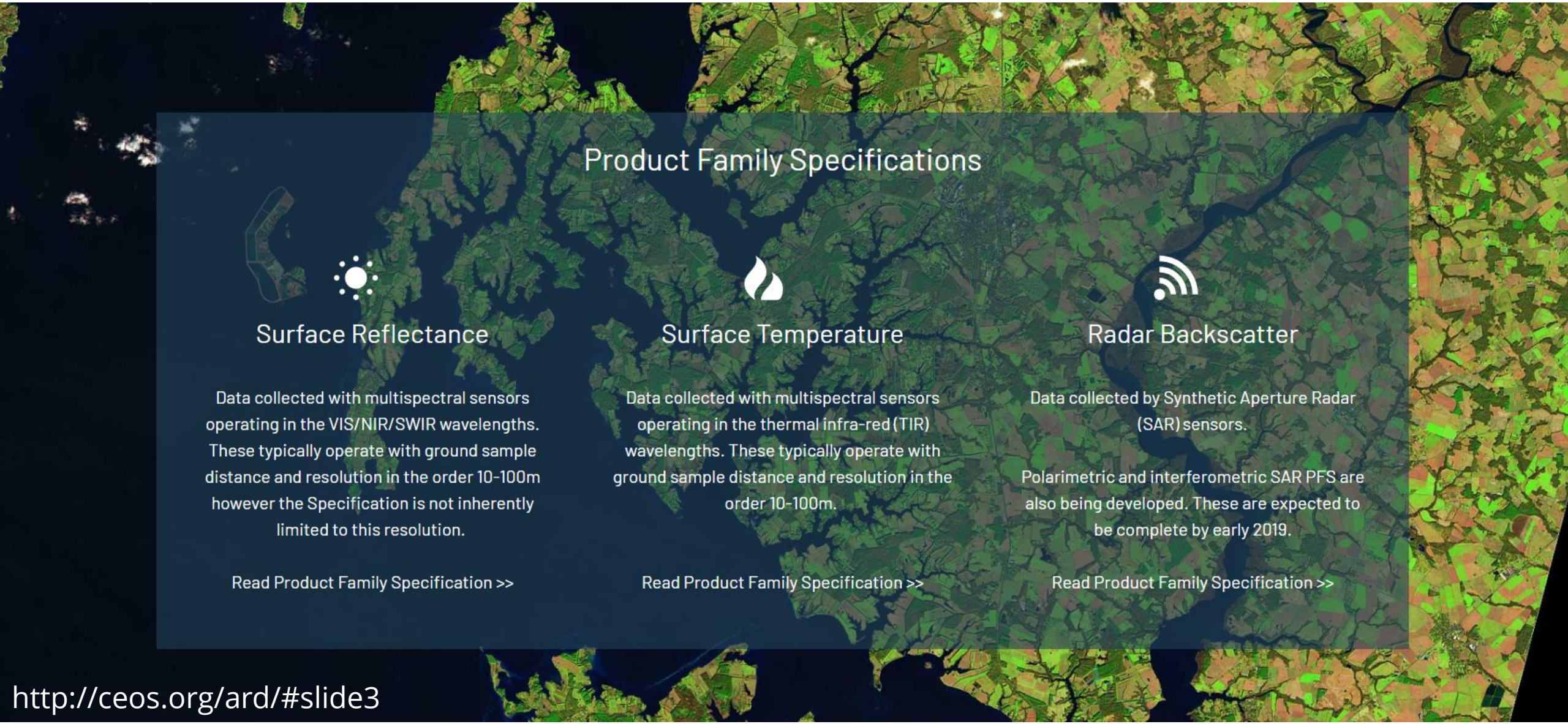
USGS

> *"U.S. Landsat Analysis Ready Data (ARD) are consistently processed to the highest scientific standards and level of processing required for direct use in monitoring and assessing landscape change."*

CEOS

> *"CEOS Analysis Ready Data for Land (CARD4L) are satellite data that have been processed to a minimum set of requirements and organized into a form that allows immediate analysis with a minimum of additional user effort and interoperability both through time and with other datasets."*

CEOS Analysis-ready data for land (CARD4L)



http://ceos.org/ard/#slide3

The FAIR principle

The FAIR Guiding Principles for scientific data management and stewardship

- **F**indable:

  Rich metadata, persistent identifier, indexed metadata.

- **A**ccessible:

  Standardised, open, free protocol for metadata access.

- **I**nteroperable:

  Dictionary, references to other metadata.

- **R**eusable:

  License, provenance, metadata quality.

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E. and Bouwman, J., 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, *3*.

Evans, Ben; Druken, Kelsey; Wang, Jingbo; Yang, Rui; Richards, Clare; Wyborn, Lesley (2017): A Data Quality Strategy to Enable FAIR, Programmatic Access across Large, Diverse Data Collections for High Performance Data Analysis. In Informatics 4 (4), p. 45. DOI: 10.3390/informatics4040045.

4.4

Information production

Concepts and technologies for big EO analytics

TAT-7 | 21.06.2019 | Martin Sudmanns, Dirk Tiede, & contrib. Z_GIS colleagues

Interfaculty Department of Geoinformatics – Z_GIS, University of Salzburg

Tasks and applications

- Indexing / finding data
- Data fusion / data enhancement
- Classification
- Detection
- Understanding
- Prediction

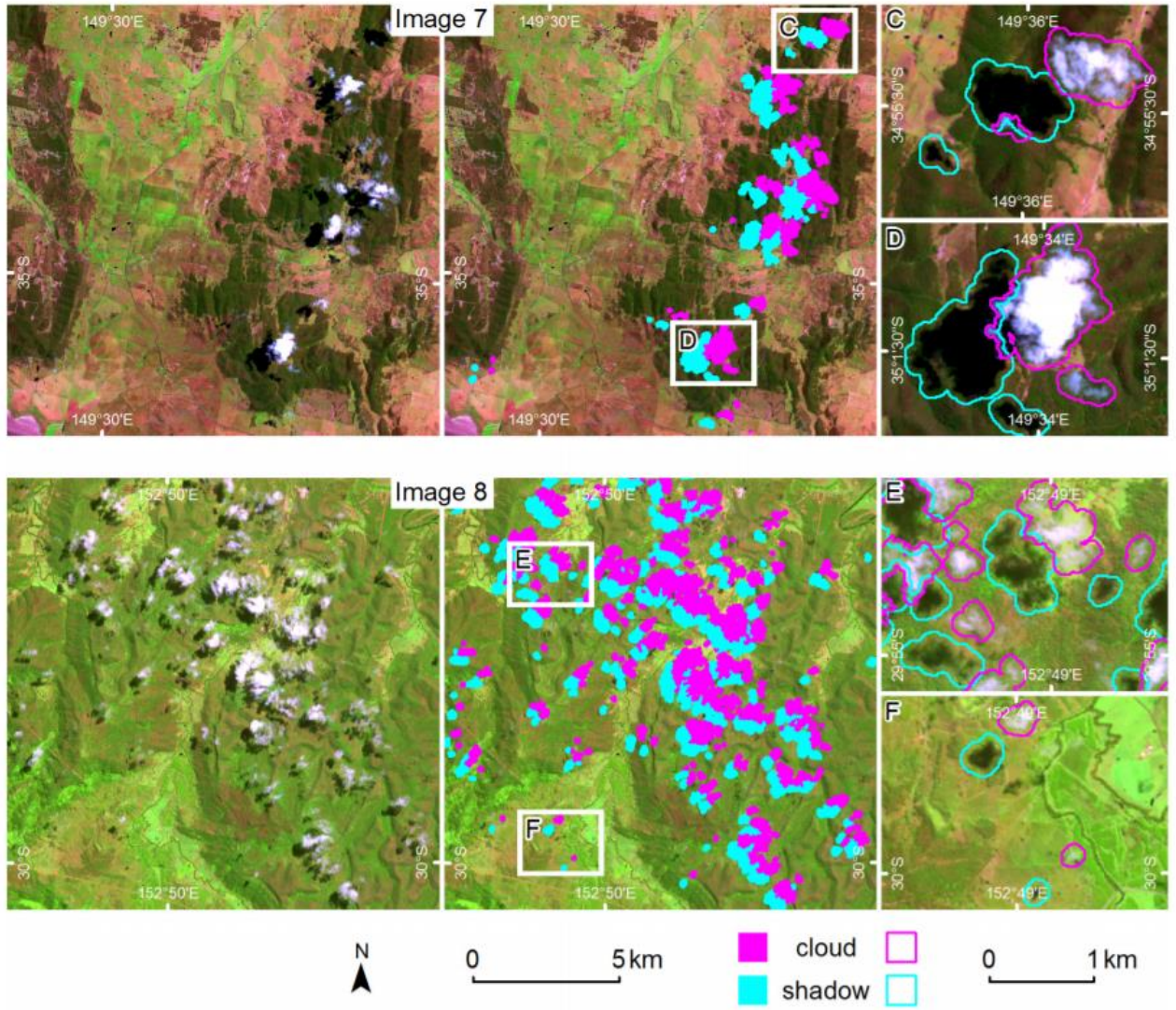
Artificial Intelligence (AI) for Earth Observation (AI4EO)

Image labelling → Object Detection → Dense labelling / semantic segmentation

https://medium.com/earthcube-stories/ai-products-and-remote-sensing-yes-it-is-hard-and-yes-you-need-a-good-infra-4b5d6cf822f1

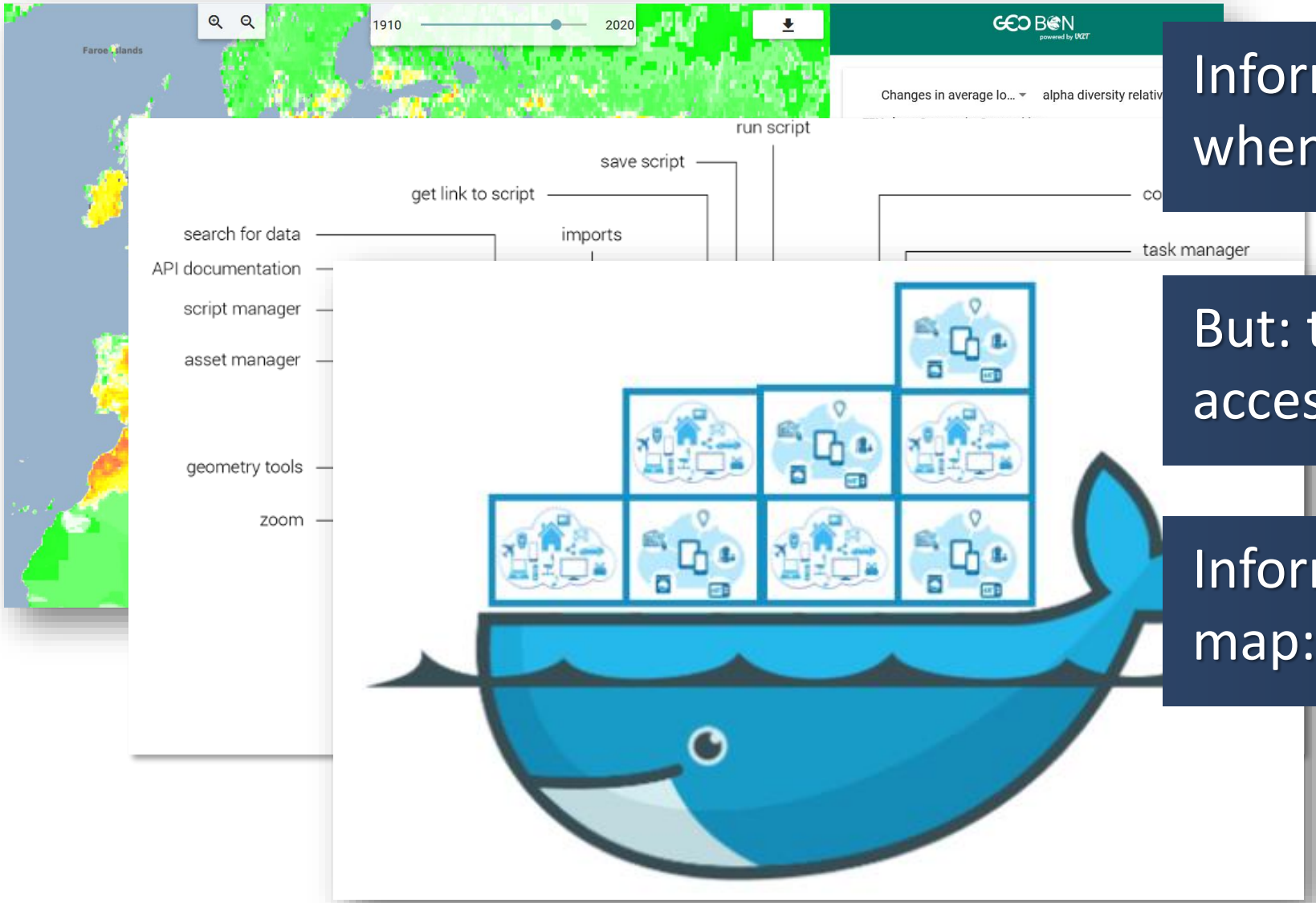
Specific nature of remote sensing data

- **Physically-based measurement** of physical quantities across the whole electromagnetic spectrum
- Many (obvious) **causal relationships**
- Variety often can‘t be captured by samples
- Examples: Cloud, shorelines, topographic shadows



Fisher, A. Cloud and Cloud-Shadow Detection in SPOT5 HRG Imagery with Automated Morphological Feature Extraction. *Remote Sens.* **2014**, *6*, 776-800.

When can information from big EO data be useful?



Information: What? Where? when? who?

But: tailored user interfaces / access

Information is more than a map: an executable container

From data to information

End users don‘t care about (big Earth) data, they want reliable information

Space 4.0: Make ***MONEY*** Big Earth data is supposed to be a business

„New kids on the block“: Data science, machine learning, …

…?

Answers

Expectations

Questions