22.05.2023

Julius Kühn-Institut
Bundesforschungsinstitut für Kulturpflanzen
Federal Research Centre for Cultivated Plants

Cloud-integrierte Geodateninfrastrukturen in der Fernerkundung

… auch in Behörden?

Florian Beyer, et al.

Heike Gerighausen, Markus Möller

FLF-Forschungszentrum für Landwirtschaftliche Fernerkundung@
Institut für Pflanzenbau und Bodenkunde

www.julius-kuehn.de

Werdegang

- * Leipzig

- **2004 – 2013 Martin-Luther-Universität Halle-Wittenberg**
  - **Diplom Geographie (NF: Bodenkunde / Geologie)**
  - **Hyperspektrale Fernerkundung in Bergbaufolgelandschaften**

- 2013 – 2017 Universität Osnabrück
  - Wissenschaftlicher Mitarbeiter
  - Landwirtschaftliche Klassifikation mit multispektralen Satellitendaten

- 2017 – 2021 Universität Rostock
  - Dr.-Ing. Geoinformatik
  - Moorvegetationsklassifikation mit multisensoralen Drohnendaten

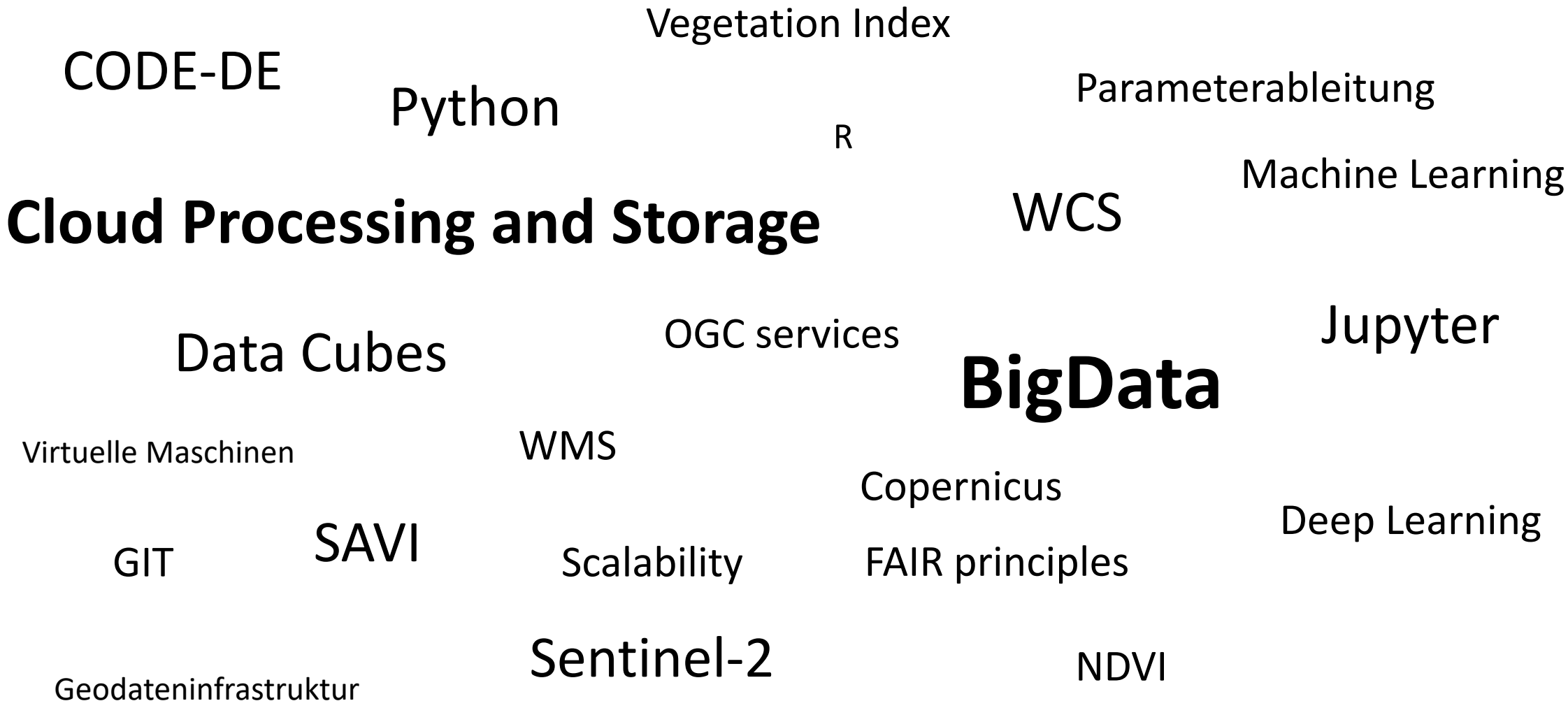
- Ab 2021 Julius-Kühn-Institut Braunschweig
  - Post-Doc
  - In-Season Crop Type Classification using Copernicus data
  - GeoDataSteward (ab 2023) FAIR-Prinzipien & FDM

Studium Uni Halle 2009

Inhalt

1. Cloud-integrierte Geodateninfrastruktur
   a. die Behördencloud CODE-DE
2. Analyse-fertige Copernicus-Daten
   a. Use Case S2_GermanyGrid
3. DataCubes
   a. Use Case Data Cubes

Buzzwords


Vegetation Index
CODE-DE
Python
Parameterableitung
R
Machine Learning
Cloud Processing and Storage
WCS
Jupyter
Data Cubes
OGC services
BigData
Virtuelle Maschinen
WMS
Copernicus
Deep Learning
SAVI
GIT
Scalability
FAIR principles
Sentinel-2
NDVI
Geodateninfrastruktur

Challenges of the 21st century

Spatial data area

callenges of 21st century

- **Climate Change**
  - Rising temperature
  - Extreme Weather events
  - Environ. degradation
- **Technology Advancements**
  - AI
  - Automation
  - Genetic engineering
- **Data Privacy and Security**
  - Protecting personal information
  - preventing data breaches
  - Responsible use of data

bigdata in geo analytics

- **Data Volume and Variety**
  - Storage
  - Management
  - Analysis
- **Data Quality and Accuracy**
  - Different sources and collection methods
  - Errors, inconsitencies, gaps
  - Requires validation and proprcessing
- **Data Integration and Interoperationality**
  - Different sources and formats
  - Interoperaibilty between different formats, standards and systems

large-scale processing

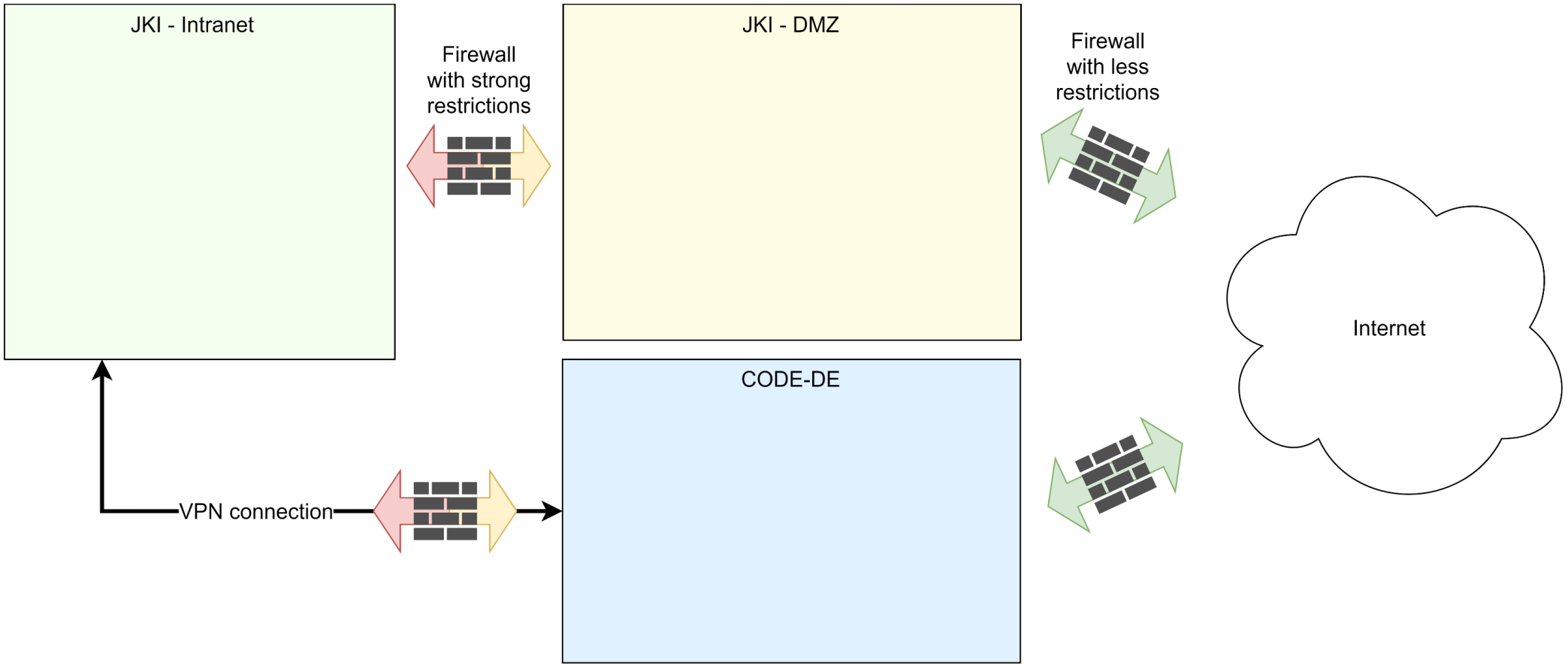
- **Scalability**
  - Scalable architectures
  - Distributed computing frameworks
- **Processing Speed**
  - Efficient algorithms and techniques
  - Parallel processing
  - Data streaming
  - Optimization techniques
- **Resource Allocation**
  - Efficient handling of processing power, storage and memory

25.01.2023

zeit.de/2021/10/buerokratie-reform-staat-corona-impfung-verwaltung
wiwo.de/my/politik/deutschland/verwaltung-wie-viel-buerokratie-braucht-deutschland/24228312.html
presseportal.de/pm/39474/5425414

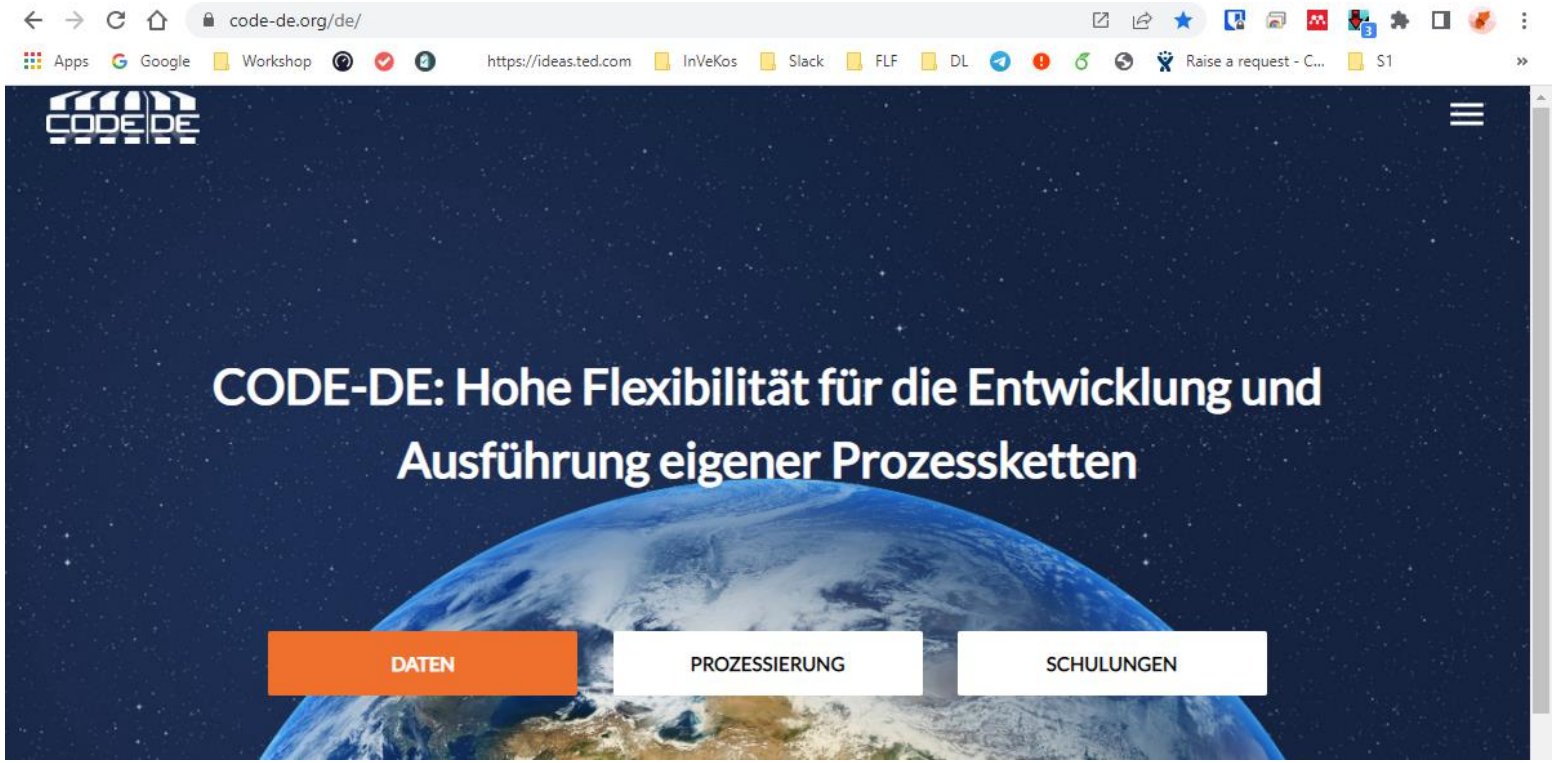
www.julius-kuehn.de

Geodateninfrastruktur des FLF



DMZ = demilitarized zone

CODE-DE – Cloud recources for public authorities



- Simple **access to Copernicus data** and **scalable cloud computing resources**
- Scalable computing power and broadband data connectivity
- Easy-to-use applications for data analysis
- **Certified by BSI**, hosted in a data center in Frankfurt (GER)
- High flexibility for the development and execution of own workflows
- Part of Germany's geoinformation strategy

FLF JKI

code-de.org

Funding

Bundesministerium für Digitales und Verkehr

Management

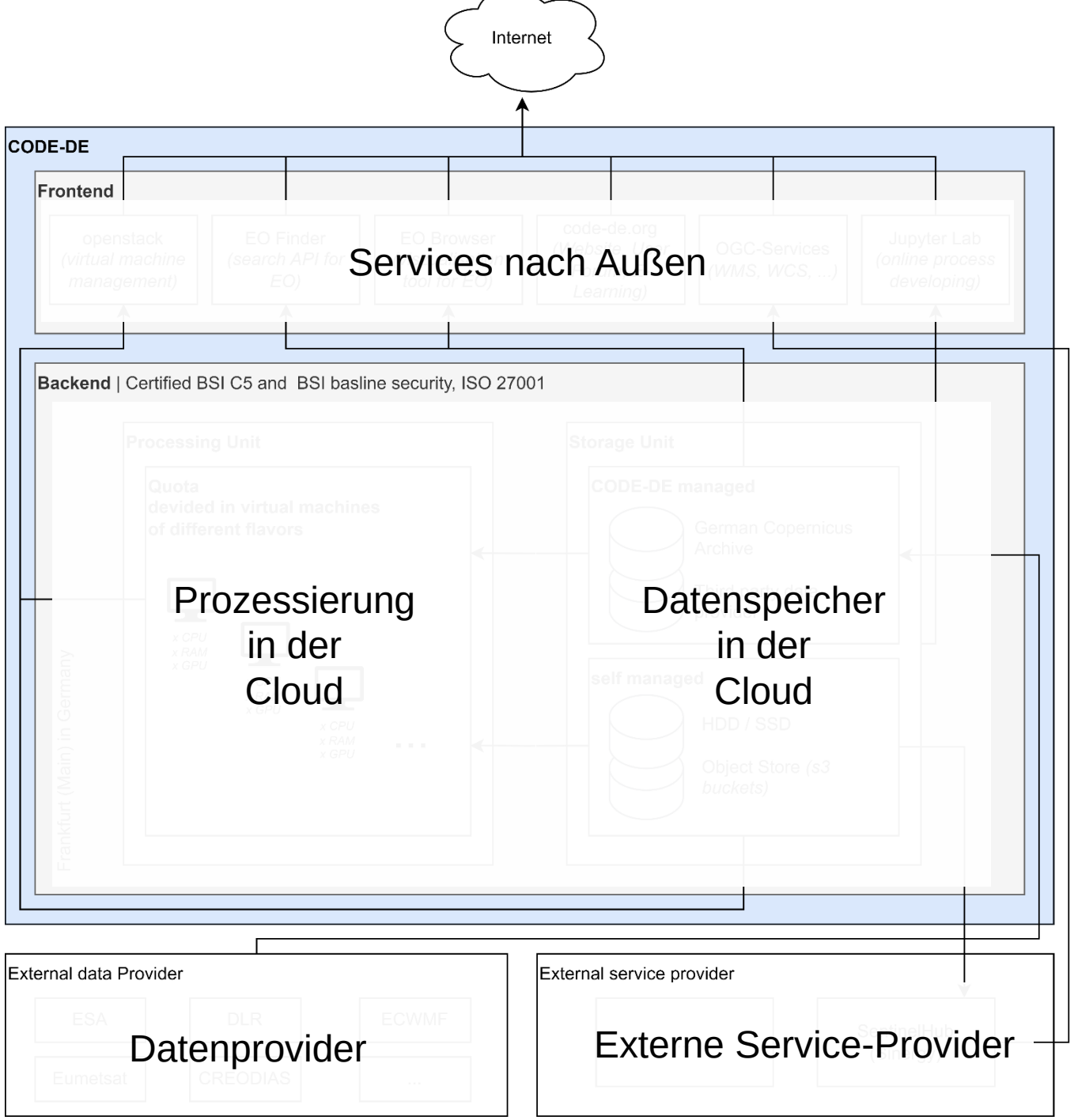
DLR

Technical realization

CloudFerro

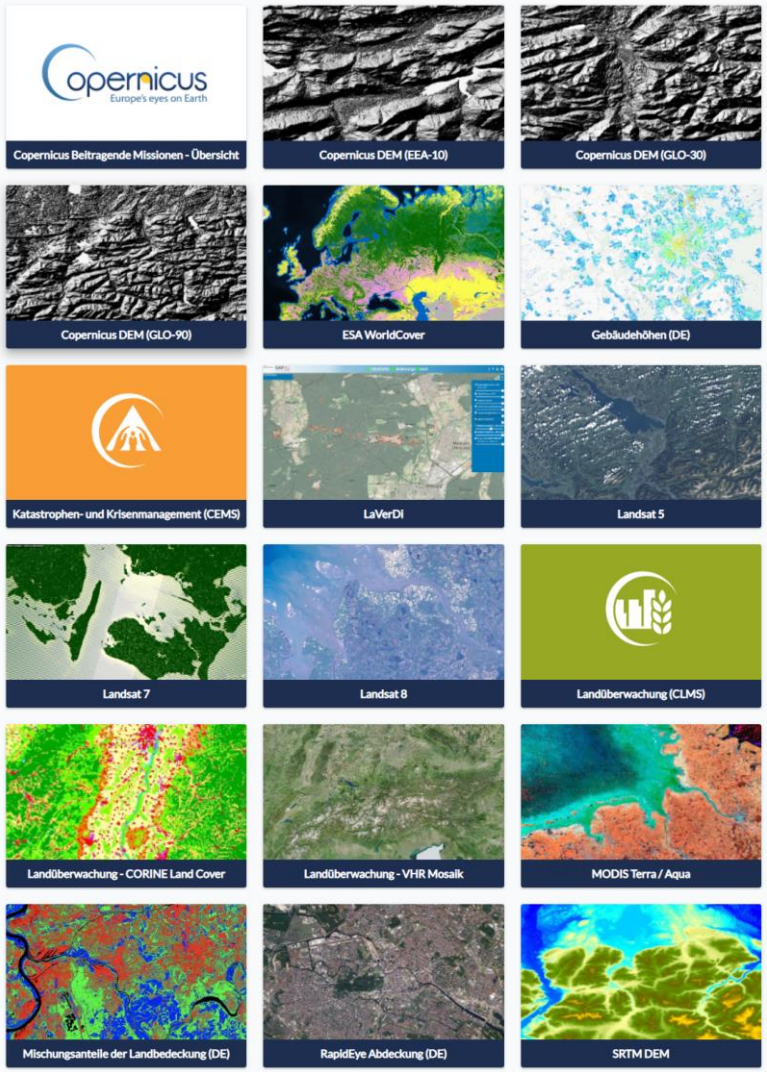
www.julius-kuehn.de

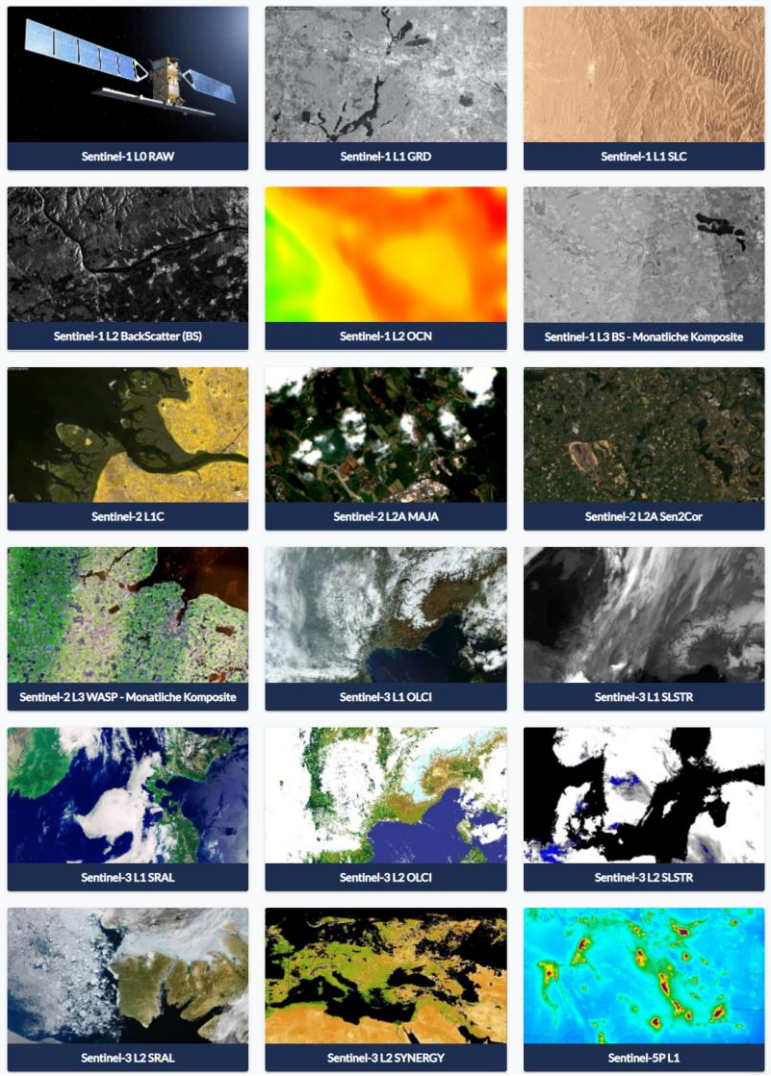
CODE-DE in a nutshell

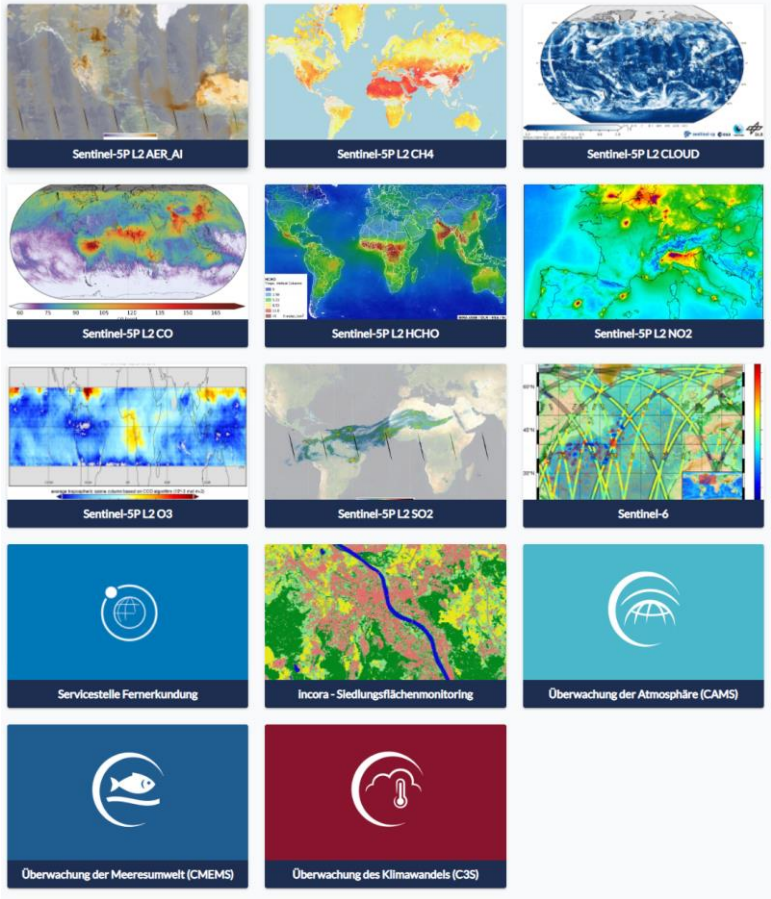


Copernicus data on CODE-DE

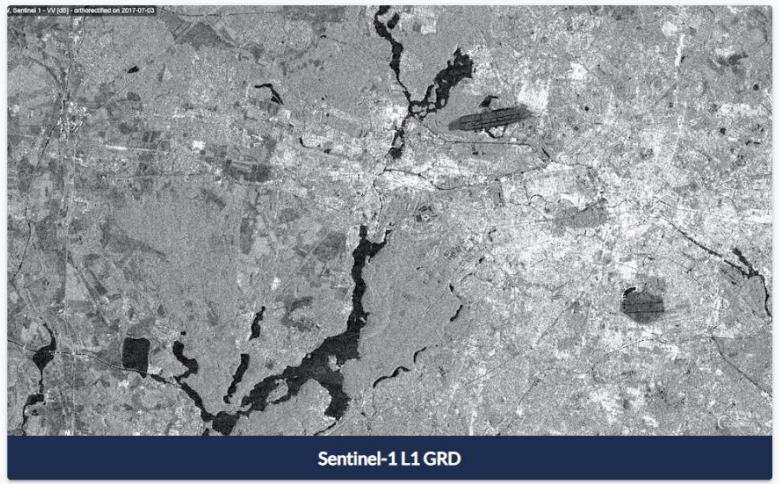
18 PB → 18.000 TB Geodaten







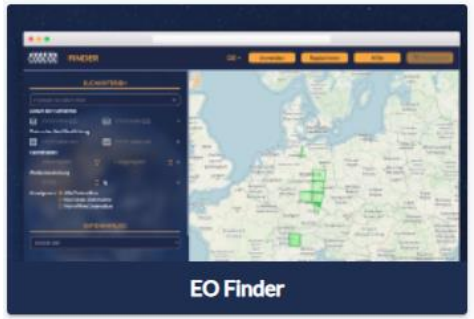
CODE-DE - Daten

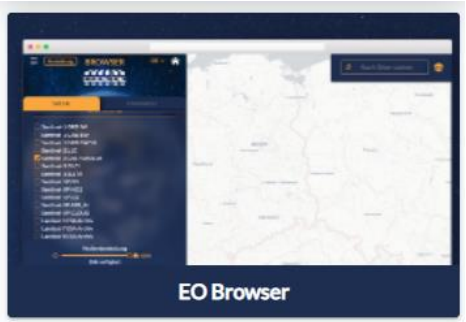

Sentinel-1 L1 GRD

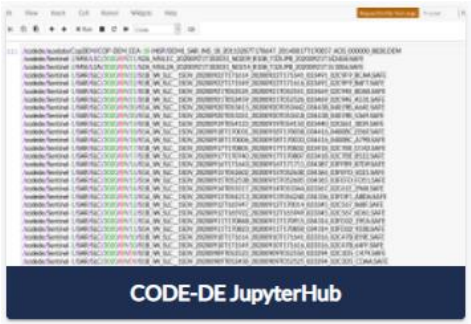

Sentinel-2 L2A Sen2Cor


Sentinel-3 L1 OLCI

CODE-DE - Prozessierung


EO Finder


EO Browser

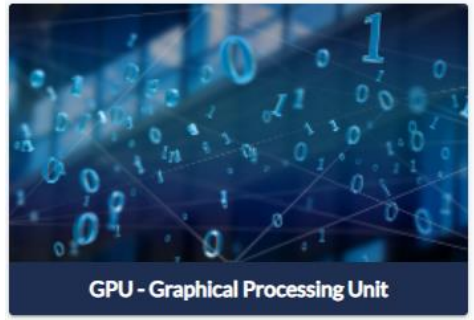

CODE-DE JupyterHub


Gehostete Prozessierung

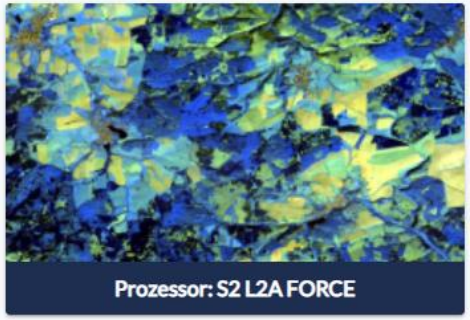

Virtuelle Maschinen

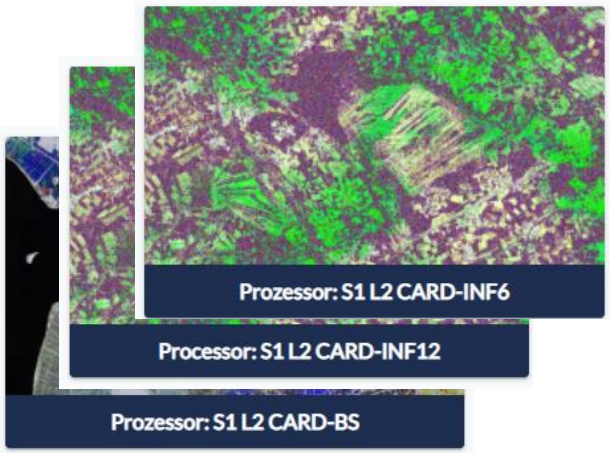

Object Storage

Volume Storage


GPU - Graphical Processing Unit

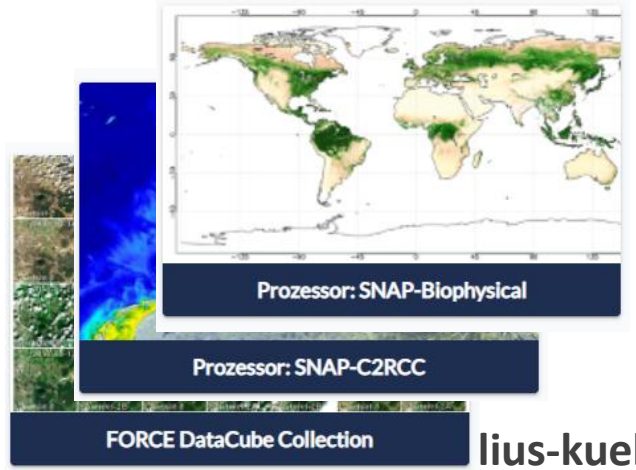

CODE-DE DataCube


Prozessor: S2 L2A FORCE


Prozessor: S1 L2 CARD-INF6

Processor: S1 L2 CARD-INF12

Prozessor: S1 L2 CARD-BS

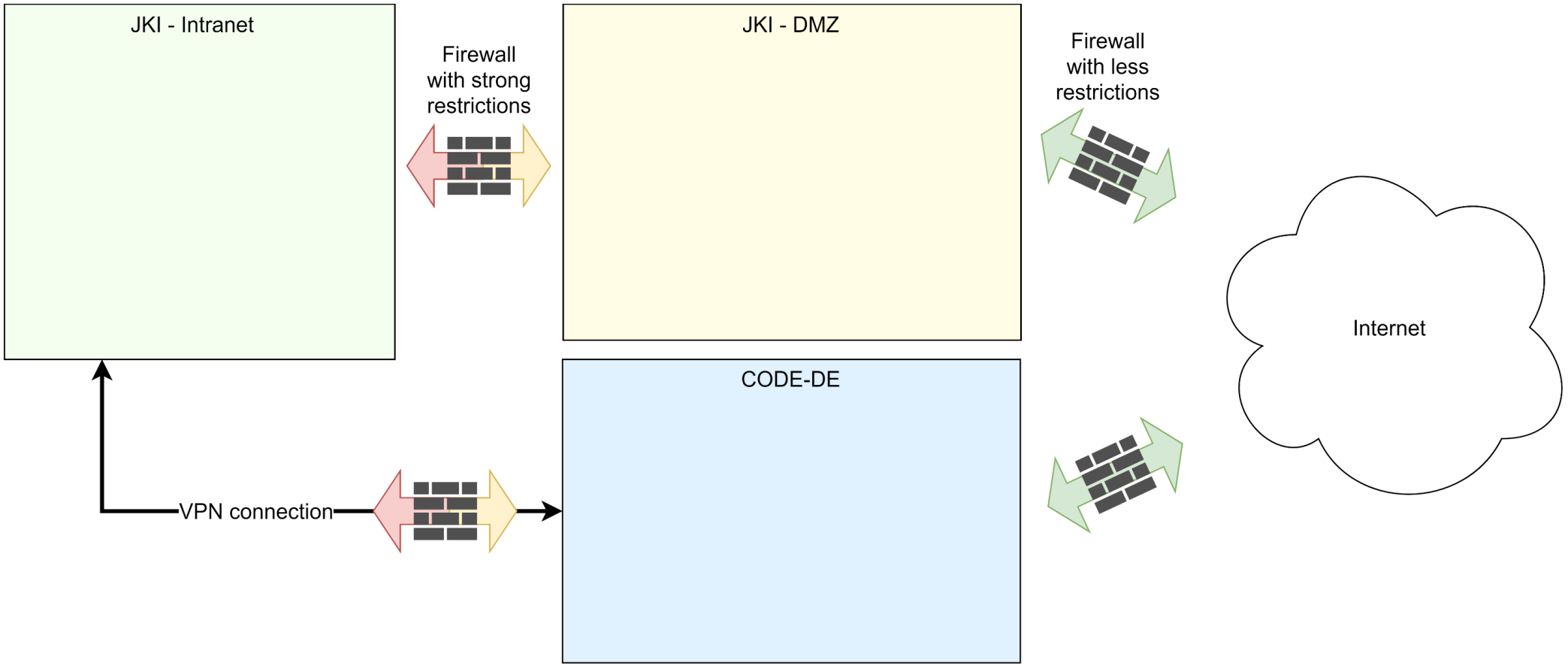

Prozessor: SNAP-Biophysical

Prozessor: SNAP-C2RCC

FORCE DataCube Collection

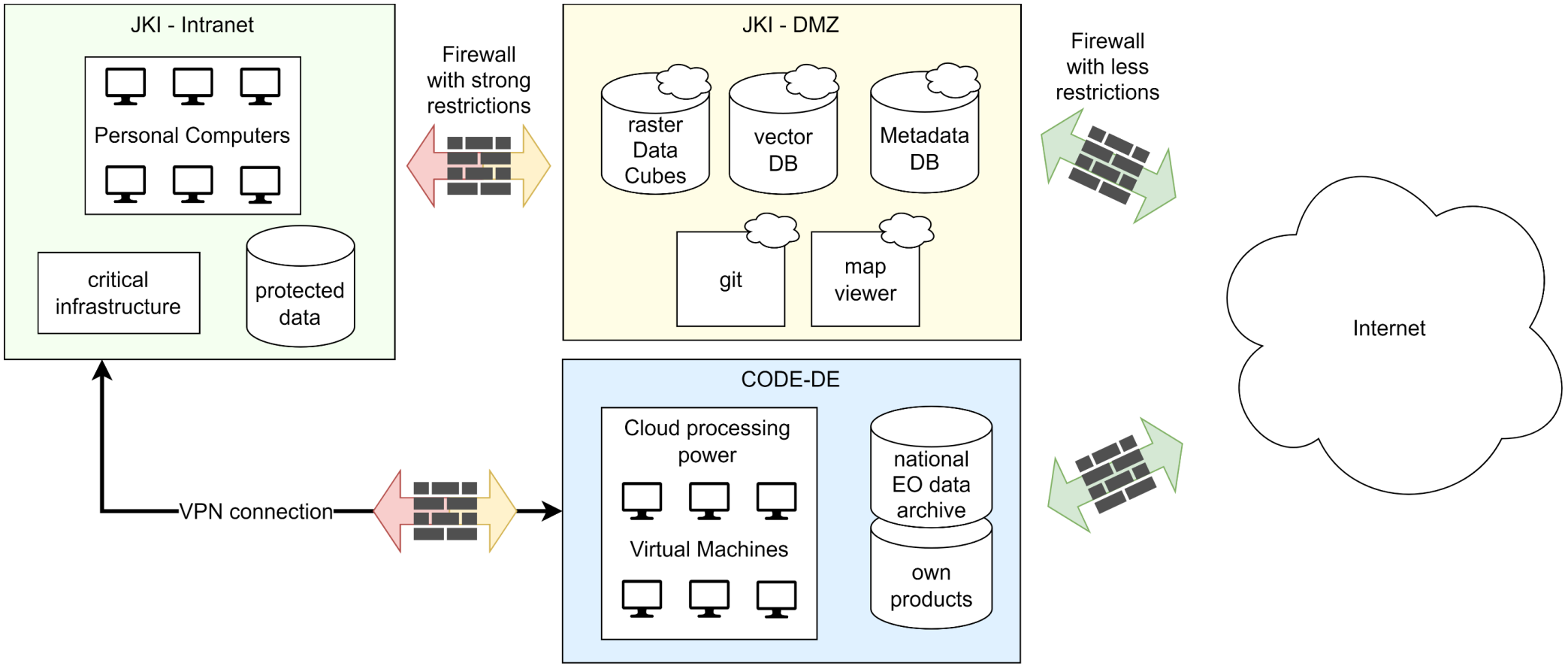
lius-kuehn.de

Geodateninfrastruktur des FLF



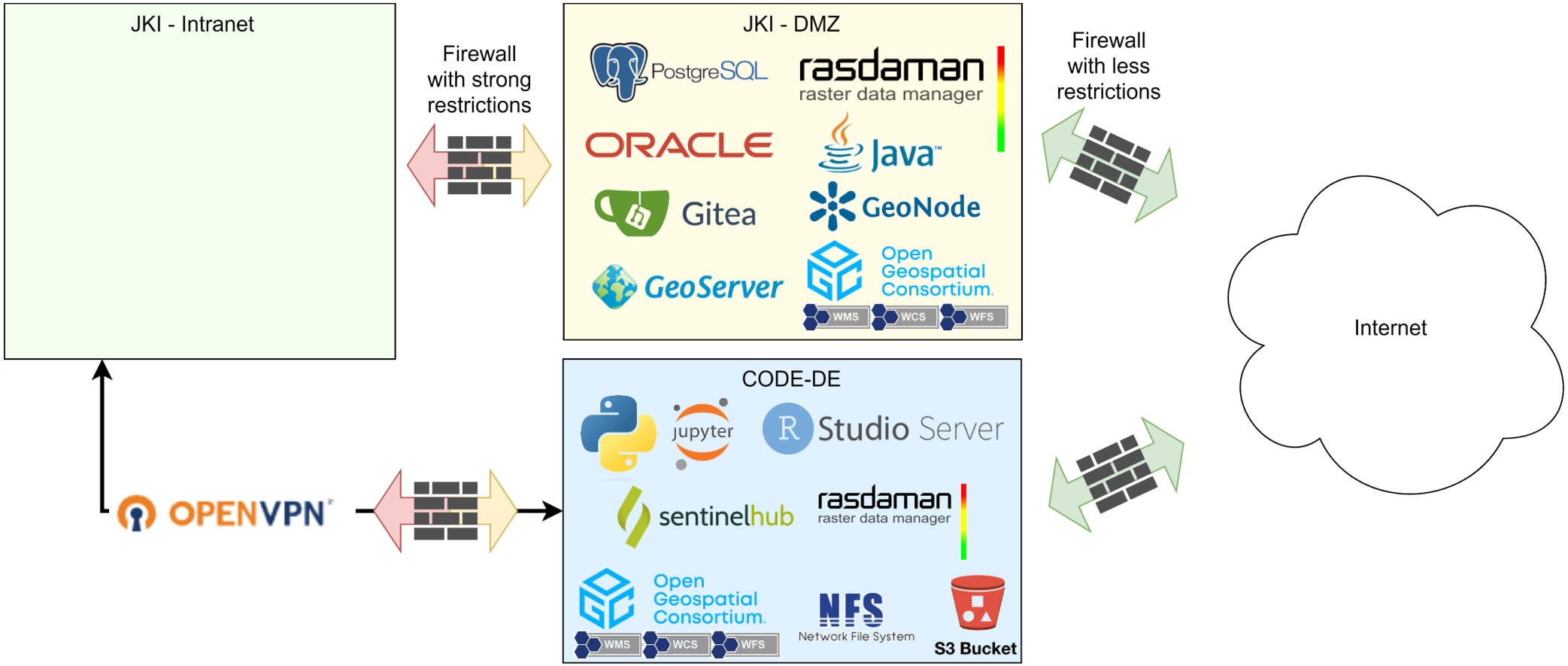
DMZ = demilitarized zone

Geodateninfrastruktur des FLF



EO = earth observation
DB = database
DMZ = demilitarized zone

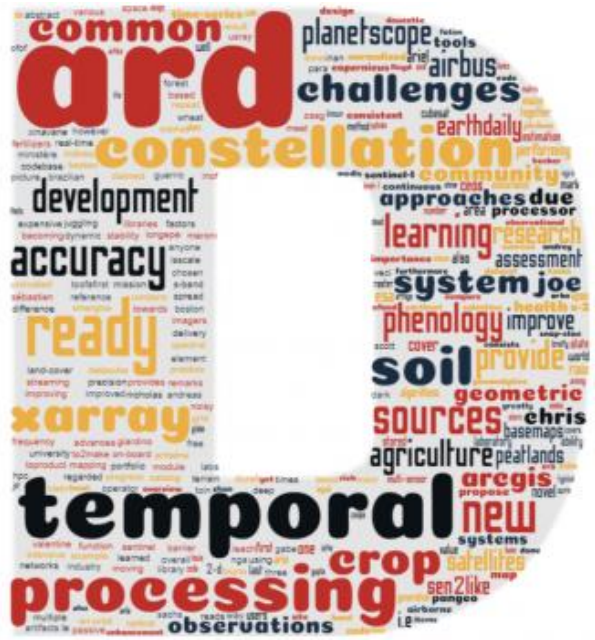
Geodateninfrastruktur des FLF

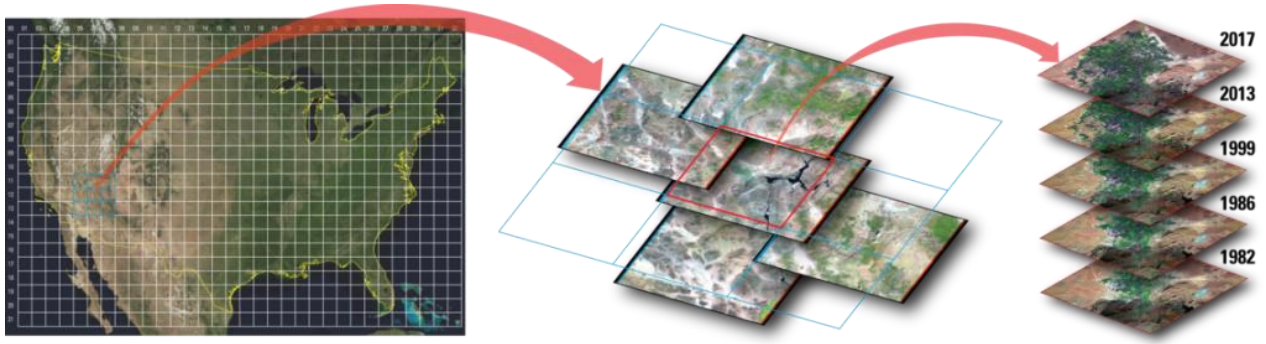


https://code-de.org/de/forum/

Allgemein → *CODE-DE-Nutzung als deutsche BEHÖRDE: Technische Lösung zur Nutzung von virtuellen Maschinen*

Analysis-ready data

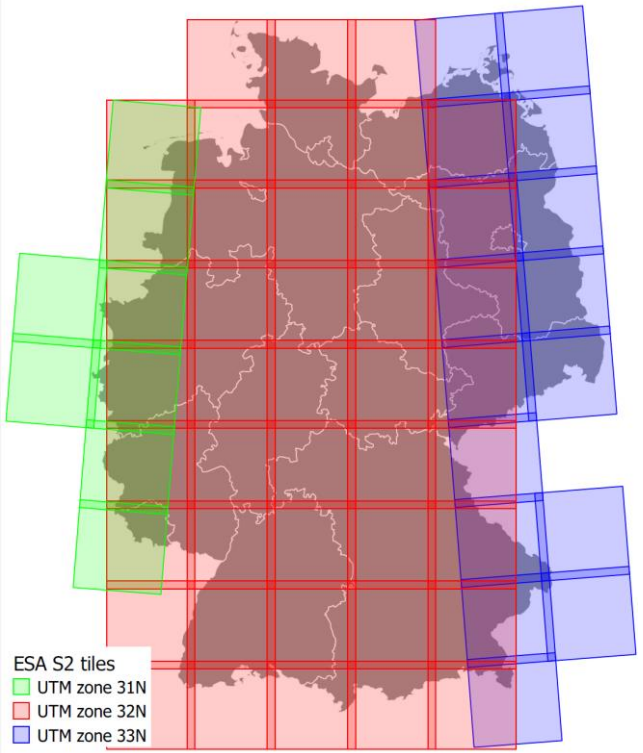




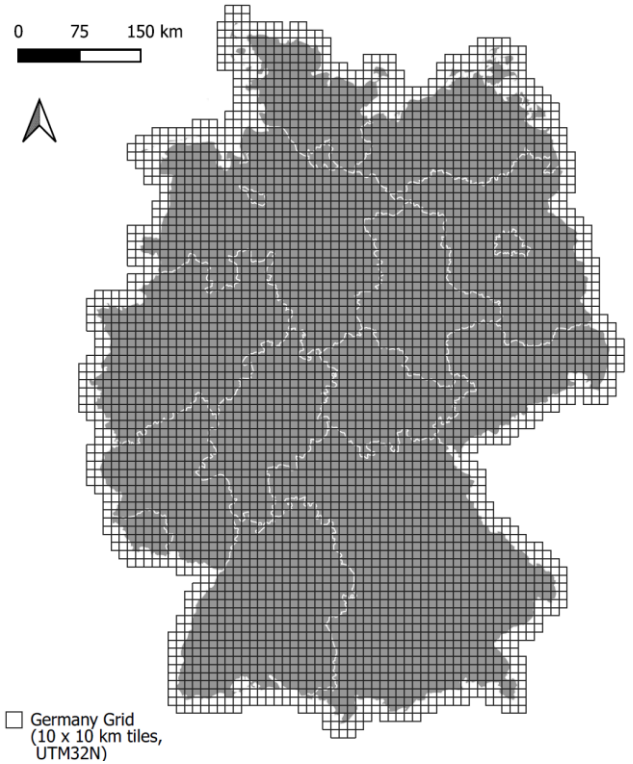
ard.zone
element84.com/blog/deconstructing-analysis-ready-data

Copernicus ARD - S2_Germany Grid

S2-L2A data from ESA

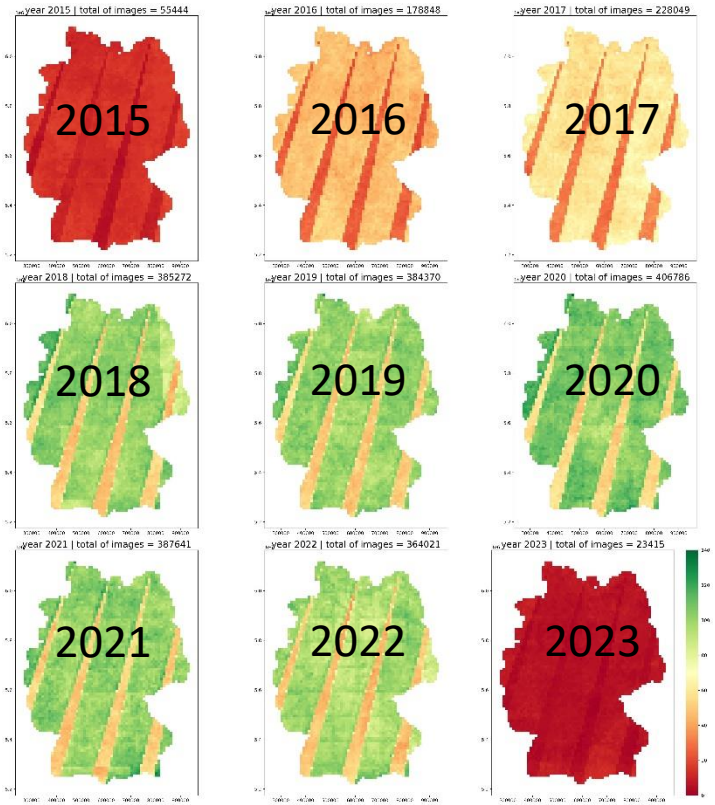


S2_Germany Grid



- UTM31 - 33/WGS84
- 10/20/60 m GSD
- 13 bands
- No masking (cloud, snow, …)
- Tiles overlap
- 0 – 10.000 %*100 Reflectance
- 16bit unit

- > 40 TB
- highly perfomant s3 online storage
- publicly available (soon)



www.julius-kuehn.de

Copernicus ARD - S2_Germany Grid

- Band selection
- Resampling
- Masking (clouds…)
- Reprojection
- Subsetting
- Stacking

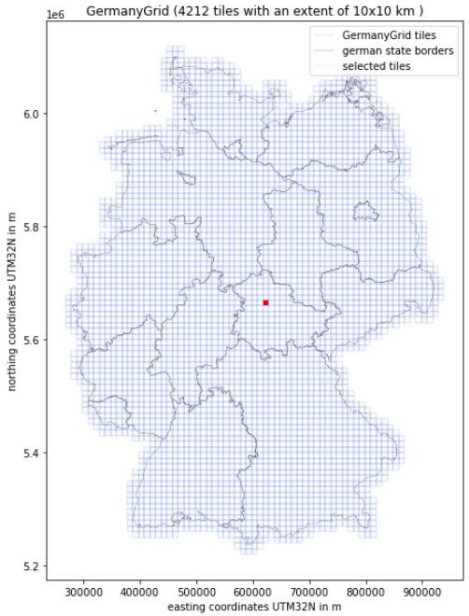
No.	Band-Sentinel-2	Center WL (nm)	Start-End WL (nm)	Band width (nm)	GSD (m)	GermanyGrid Bands	GermanyGrid Band names	GSD (M)
1	Coastalaerosol	442	433-453	20	60			
2	Blue	492	458-523	65	10	1	B	10
3	Green	559	543-578	35	10	2	G	10
4	Red	664	650-680	30	10	3	R	10
5	VegetationRedEdge	704	698-713	15	20	4	RE1	10
6	VegetationRedEdge	740	733-748	15	20	5	RE2	10
7	VegetationRedEdge	782	773-793	20	20	6	RE3	10
8	Near-Infrared	832	785-900	115	10	7	NIR10	10
8a	NarrowNIR	864	855-875	20	20	8	NIR20	10
9	Watervapor	945	935-955	20	60			
10	Shortwaveinfrared-Cirrus	1373	1360-1390	30	60			
11	Shortwaveinfrared	1610	1565-1655	90	20	9	SWIR1	10
12	Shortwaveinfrared	2200	2100-2280	180	20	10	SWIR20	10

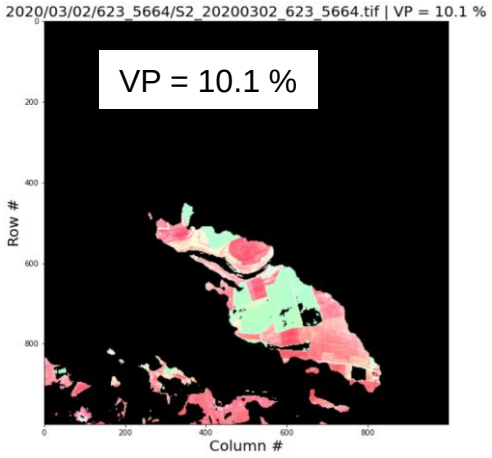
https://gitea.julius-kuehn.de/FLF/pyGERMANY_GRID.git

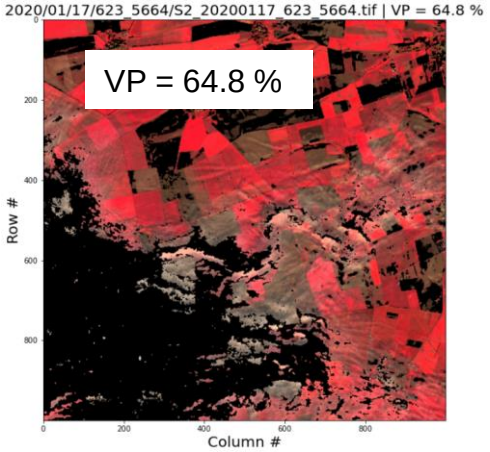
Copernicus ARD - S2_Germany Grid

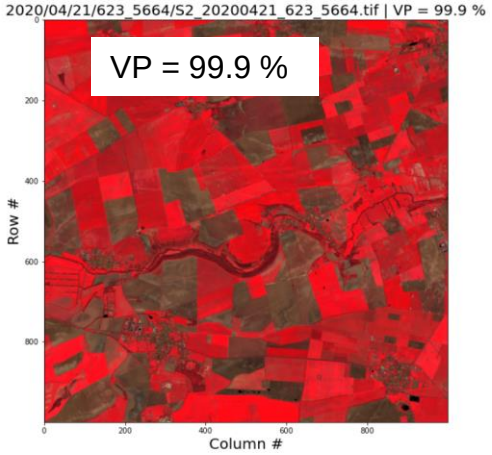
- Band selection
- Resampling
- Masking (clouds…)
- Reprojection
- Subsetting
- Stacking
  - 4,212 tiles
  - 10 x 10 km (=1,000 x 1,000 Pixel)
  - UTM32/WGS84
  - 10 m GSD
  - 10 bands (2,3,4,5,6,7,8,8a,11,12)
  - Masked: cloud, cloud shadow, snow, …
  - No overlapping
  - 0 – 10,000 %*100 Reflectance
  - 16bit unit

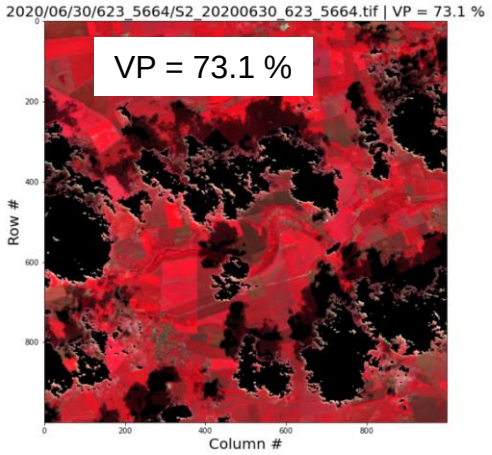
  - VP = Valid Pixel portion in %
    - stored in tiff meta data

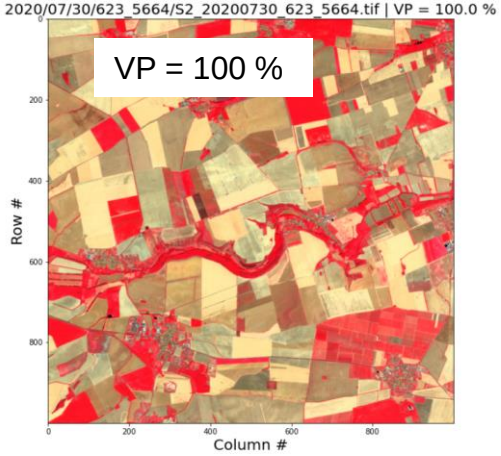






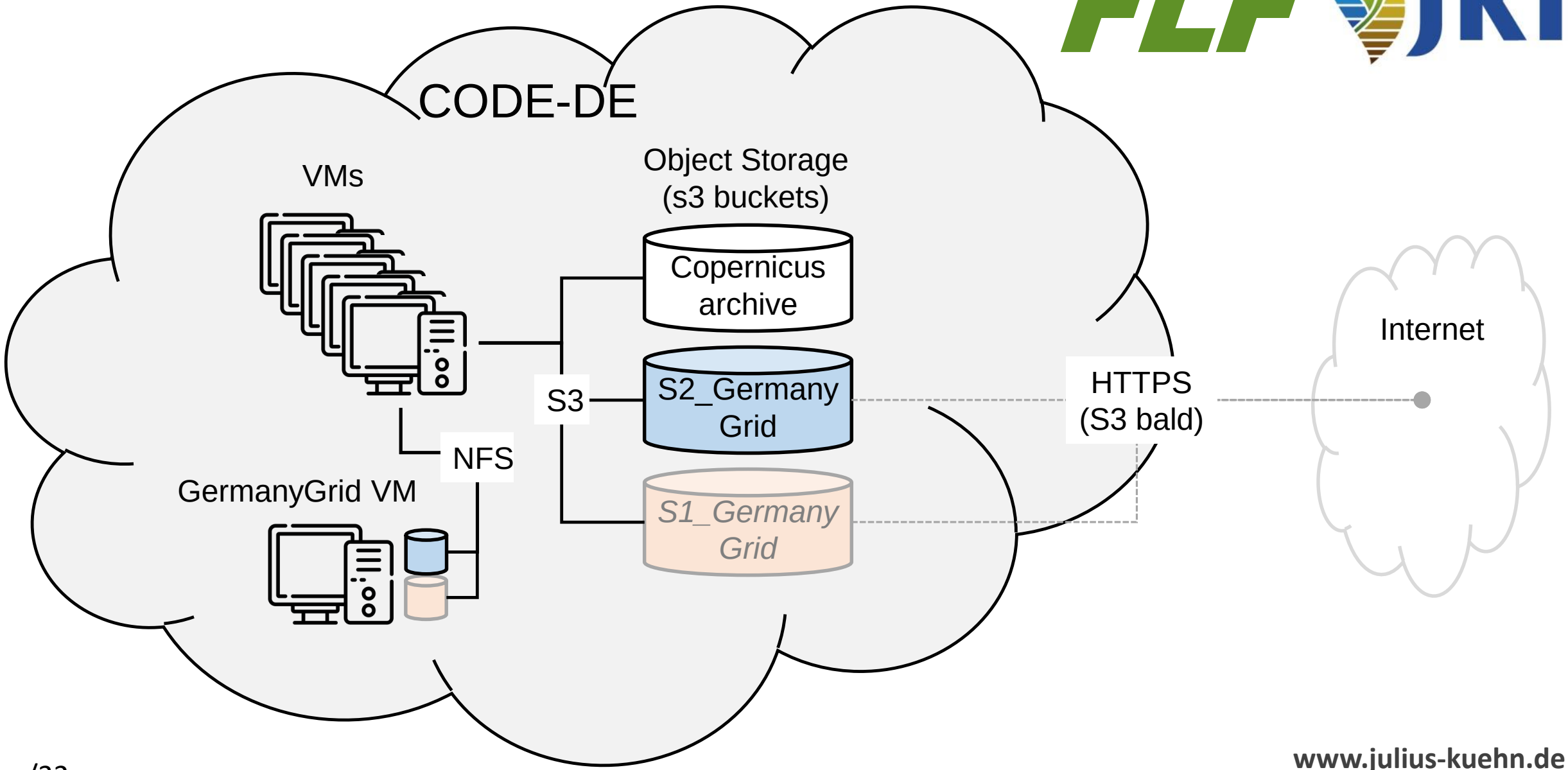






https://gitea.julius-kuehn.de/FLF/pyGERMANY_GRID.git

Copernicus ARD - S2_Germany Grid



S2_Germany Grid structure

Images stored as:

filename

S2_GermanyGrid/Year/Month/Day/TileCode/**S2_YYYYMMDD_TileCode.tif**

bucket name

key

Lists with image overview:

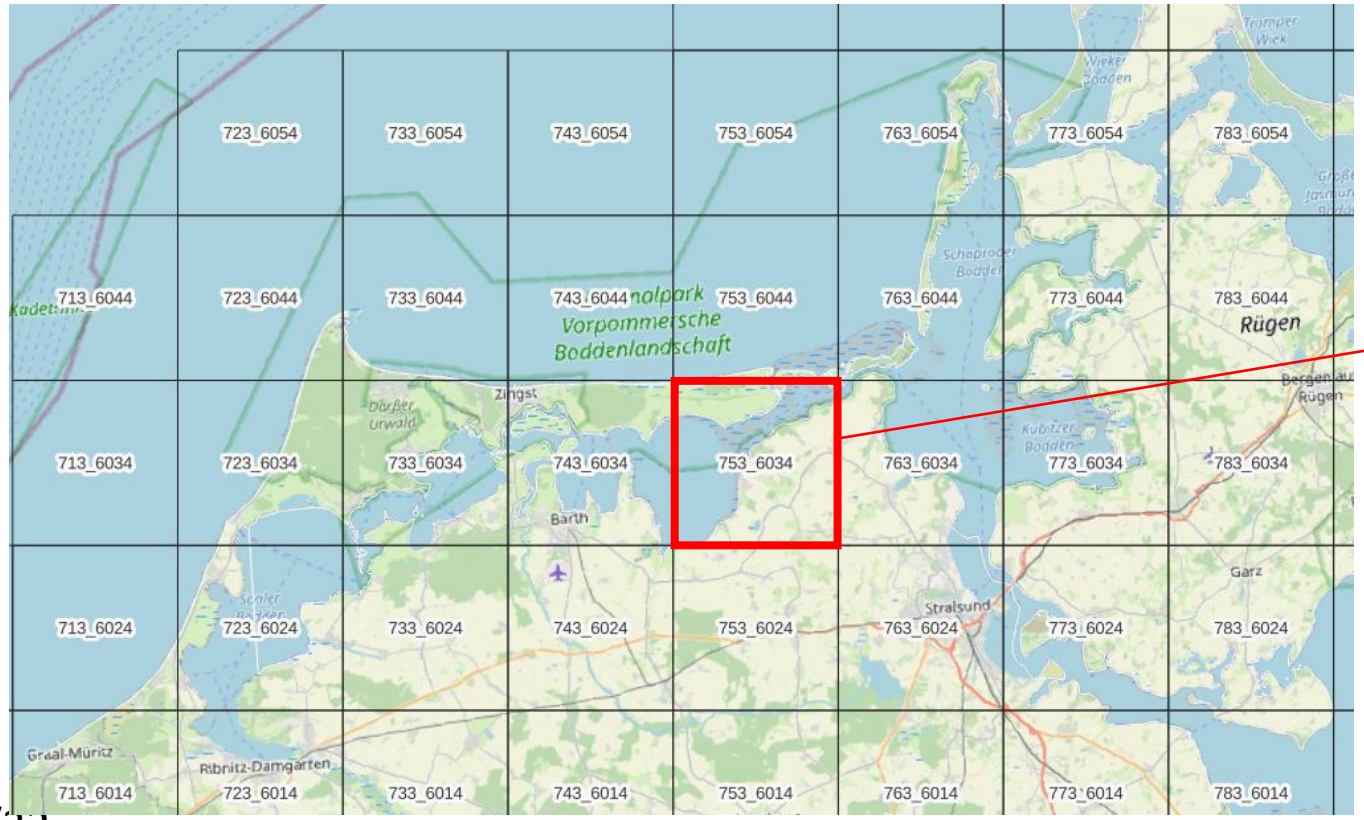
S2_GermanyGrid/Year/**YYYY_GerGrid_tiles.pkl**
S2_GermanyGrid/Year/**YYYY_GerGrid_tiles.csv**

```
'2020/01/01/273_5624/S2_20200101_273_5624.tif',
'2020/01/01/273_5634/S2_20200101_273_5634.tif',
'2020/01/01/273_5644/S2_20200101_273_5644.tif',
'2020/01/01/273_5654/S2_20200101_273_5654.tif',
'2020/01/01/273_5664/S2_20200101_273_5664.tif',
'2020/01/01/273_5674/S2_20200101_273_5674.tif',
'2020/01/01/283_5614/S2_20200101_283_5614.tif',
'2020/01/01/283_5624/S2_20200101_283_5624.tif',
'2020/01/01/283_5634/S2_20200101_283_5634.tif',
'2020/01/01/283_5644/S2_20200101_283_5644.tif',
'2020/01/01/283_5654/S2_20200101_283_5654.tif',
'2020/01/01/283_5664/S2_20200101_283_5664.tif',
'2020/01/01/283_5674/S2_20200101_283_5674.tif',
'2020/01/01/283_5684/S2_20200101_283_5684.tif',
'2020/01/01/283_5724/S2_20200101_283_5724.tif',
'2020/01/01/283_5734/S2_20200101_283_5734.tif',
'2020/01/01/283_5744/S2_20200101_283_5744.tif',
'2020/01/01/283_5754/S2_20200101_283_5754.tif',
'2020/01/01/293_5614/S2_20200101_293_5614.tif',
'2020/01/01/293_5624/S2_20200101_293_5624.tif',
'2020/01/01/293_5634/S2_20200101_293_5634.tif',
'2020/01/01/293_5644/S2_20200101_293_5644.tif',
'2020/01/01/293_5654/S2_20200101_293_5654.tif'
```

S2_Germany Grid tile code

Images stored as:

S2_GermanyGrid/Year/Month/Day/TileCode/**S2_YYYYMMDD_TileCode.tif**



Tile Code:
753_6034

Centroid coordinate:
Long: **753‘**033 m
Lat: **6‘034‘**369 m

S2_Germany Grid

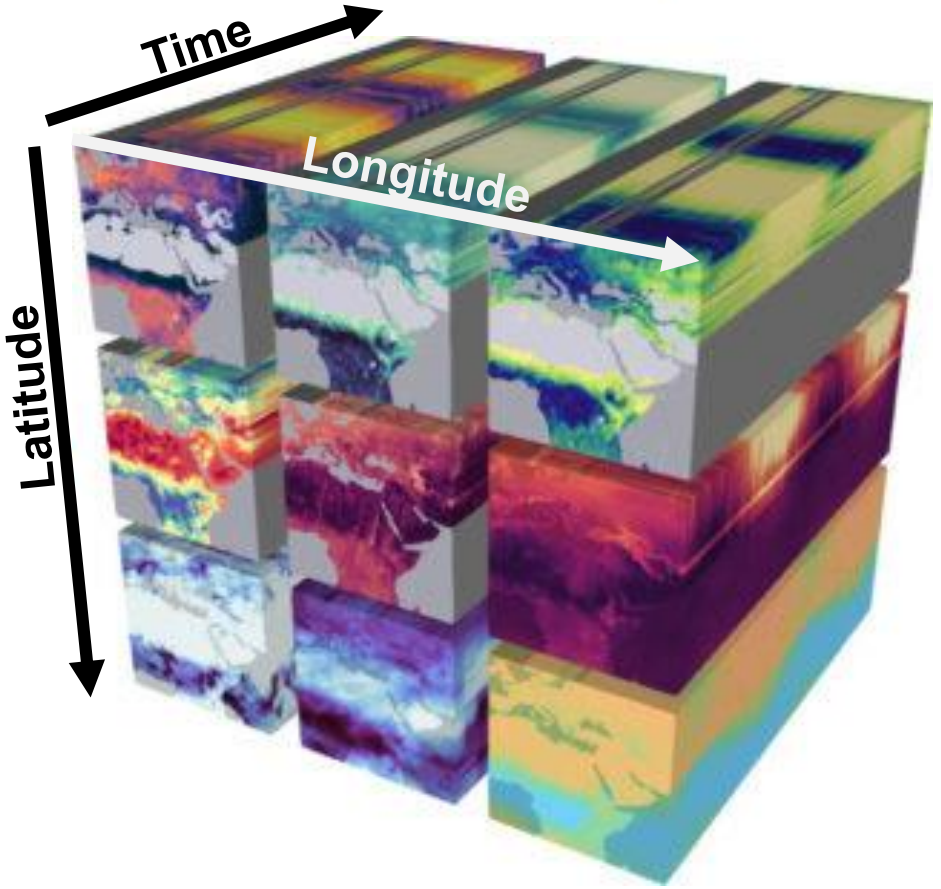
https://cloud.fra1-1.cloudferro.com:8080/swift/v1/AUTH_63f2e3111c6d4f61924ab4f410af382c/S2_GermanyGrid/

DEMO

S2_GermanyGrid

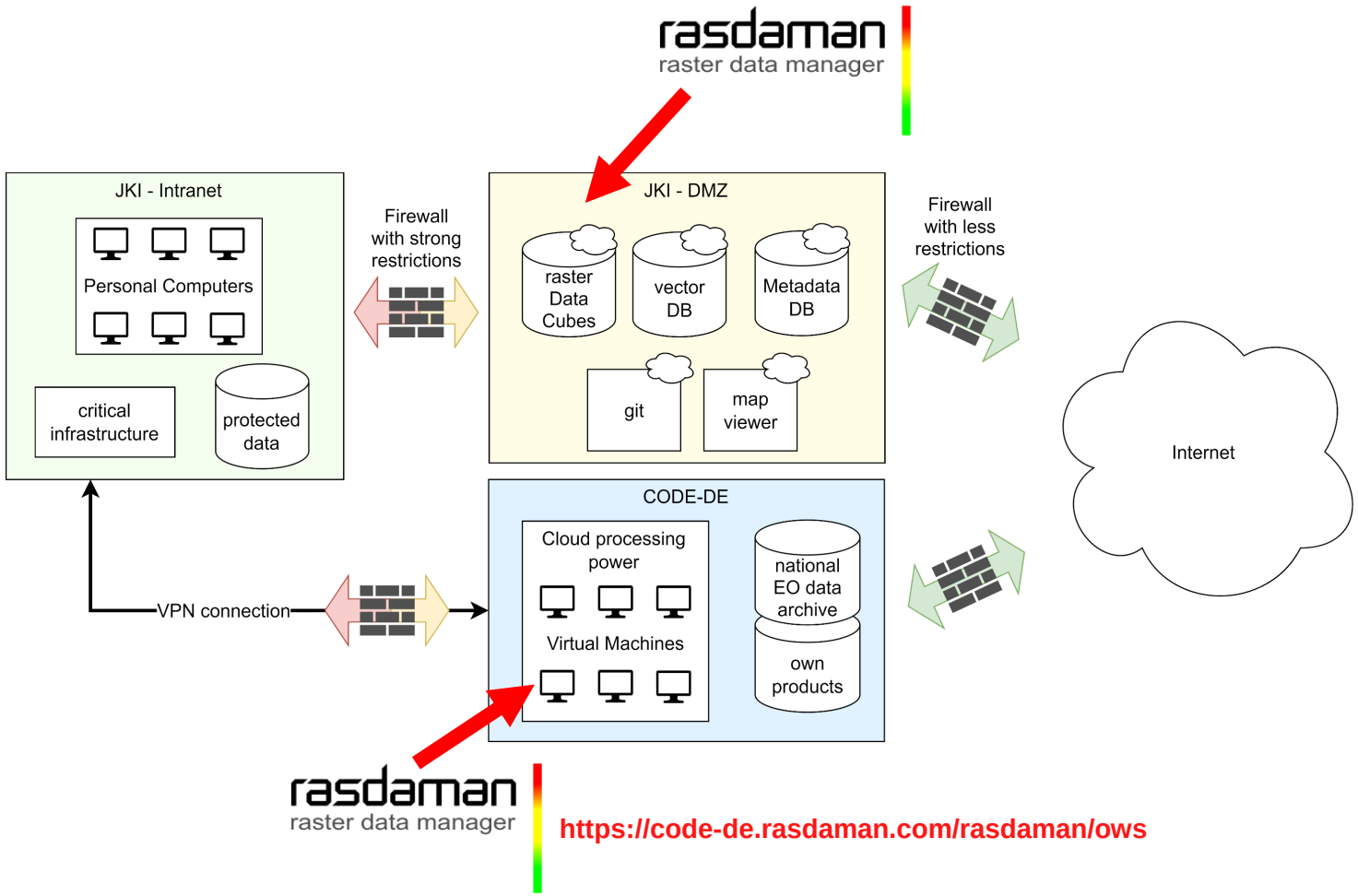
Data Cubes

Cutted, SOURCE: esa.int/ESA_Multimedia/Videos/2018/07/Earth_System_Data_Lab



Mahecha et al. : Earth system data cubes unravel global multivariate dynamics, Earth Syst. Dynam., 11, 201–234, https://doi.org/10.5194/esd-11-201-2020, 2020.

Geodateninfrastruktur des FLF



EO = earth observation
DB = database
DMZ = demilitarized zone

Exampe WCS Query

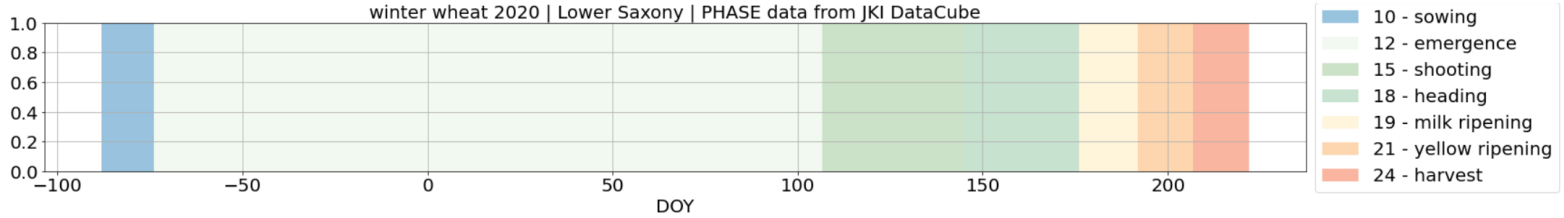
https://code-de.rasdaman.com/rasdaman/ows
?&***SERVICE***=WCS
&***VERSION***=2.0.1
&***REQUEST***=GetCoverage
&***COVERAGEID***=**S2_GermanyGrid**
&***SUBSET***=ansi("2019-10-05")
&***subsettingCrs***=http://ows.rasdaman.org/def/crs/EPSG/0/32632
&***CLIP=POLYGON***((595117.1423555185 5784603.23154566, 595123.4425648759 5784792.817361793, 595123.4467946623 5784792.959964871, 595131.3773480996 5785096.90776269, 595137.0802208507 5785261.194424775, 595196.9548982648 5785262.269346152, 595198.8303912097 5785262.391339741, 595200.6861515986 5785262.688845372, 595202.5057914145 5785263.159235806, 595204.2732416177 5785263.798357079, 595205.9728940486 5785264.600565179, 595206.692532367 5785265.02705322, 595401.684824324 5785269.974576705, 595446.0108982538 5785271.067448465, 595446.2380262674 5785270.252888932, 595455.45978537 5785242.587512609, 595499.7201932578 5785095.77155718, 595554.7875947909 5784905.712404071, 595615.3133702298 5784693.660278281, 595615.3225543081 5784693.628203245, 595653.0884108037 5784562.146929435, 595126.916582937 5784537.355768103, 595118.9777214284 5784577.541293625, 595117.192846779 5784601.068756027, 595117.1778552582 5784602.893008772, 595117.1423555185 5784603.23154566))
&***outputCrs***=http://ows.rasdaman.org/def/crs/EPSG/0/32632
&***FORMAT***=image/tiff
&***RANGESUBSET***=NIR10,R,G

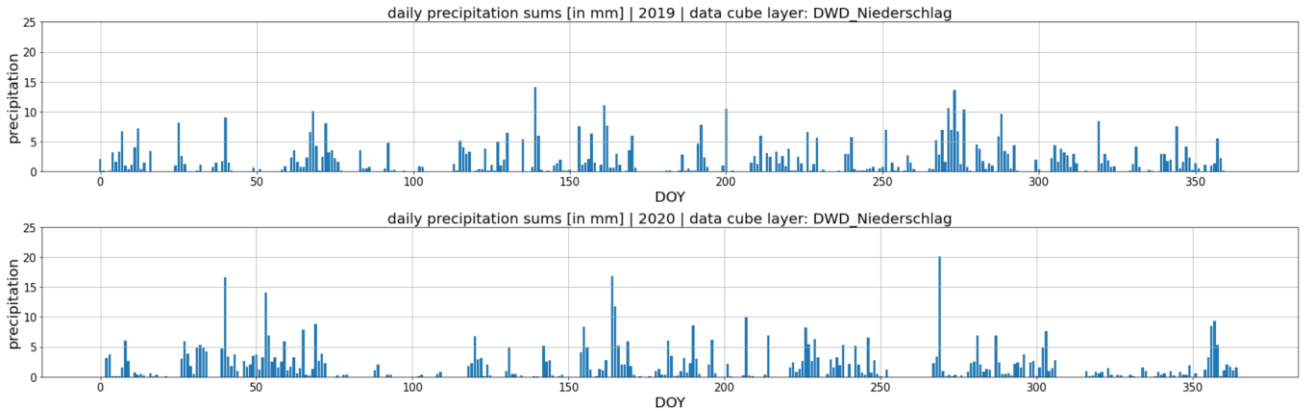
Credentials:

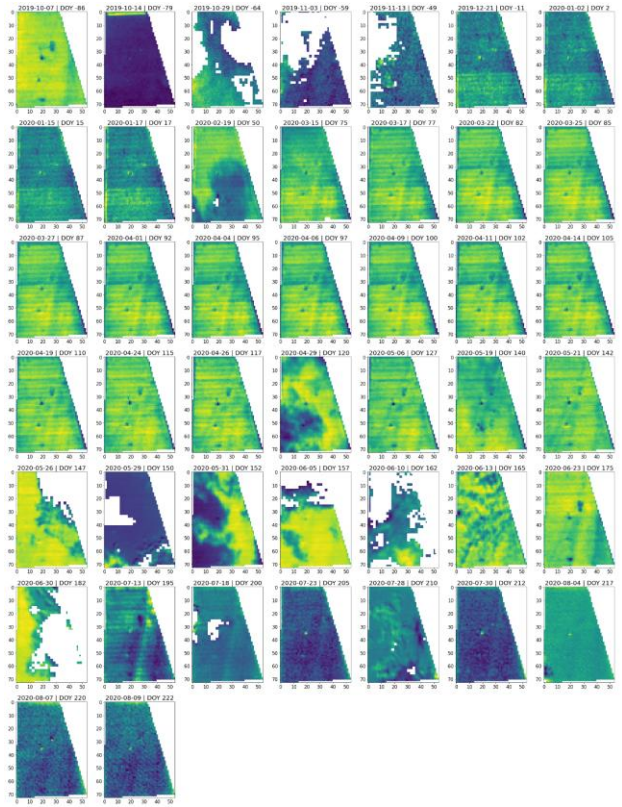
user=,***'

password=,***'

Demo Data Cubes







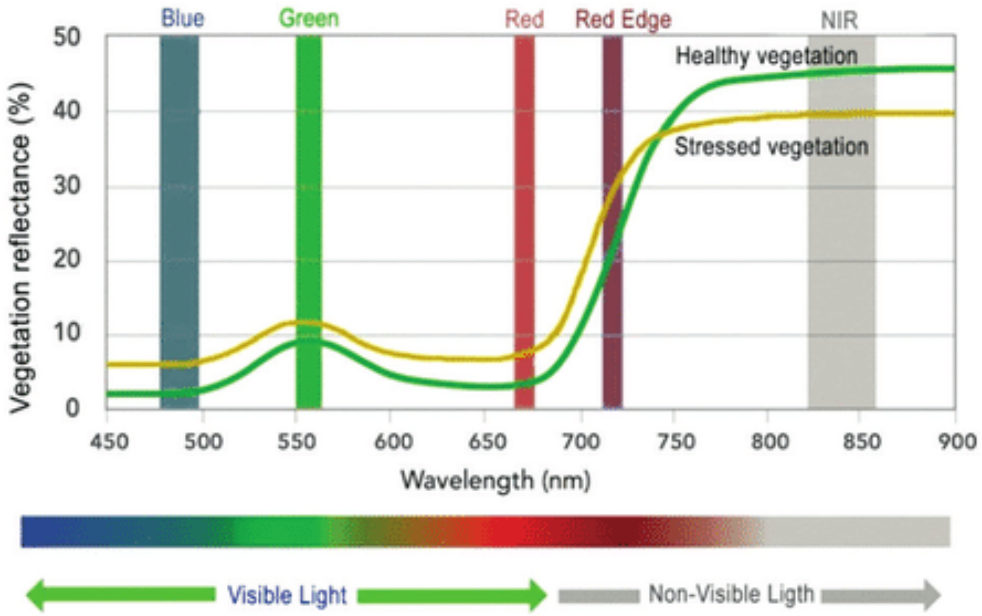
Demo Data Cubes

SAVI - Soil Adjusted Vegetation Index (indexdatabase.de)

$$\frac{(R_{800nm} - R_{670nm})}{(R_{800nm} + R_{670nm} + L)}(1 + L)$$

$$-0.9 \leq L \geq 1.6$$

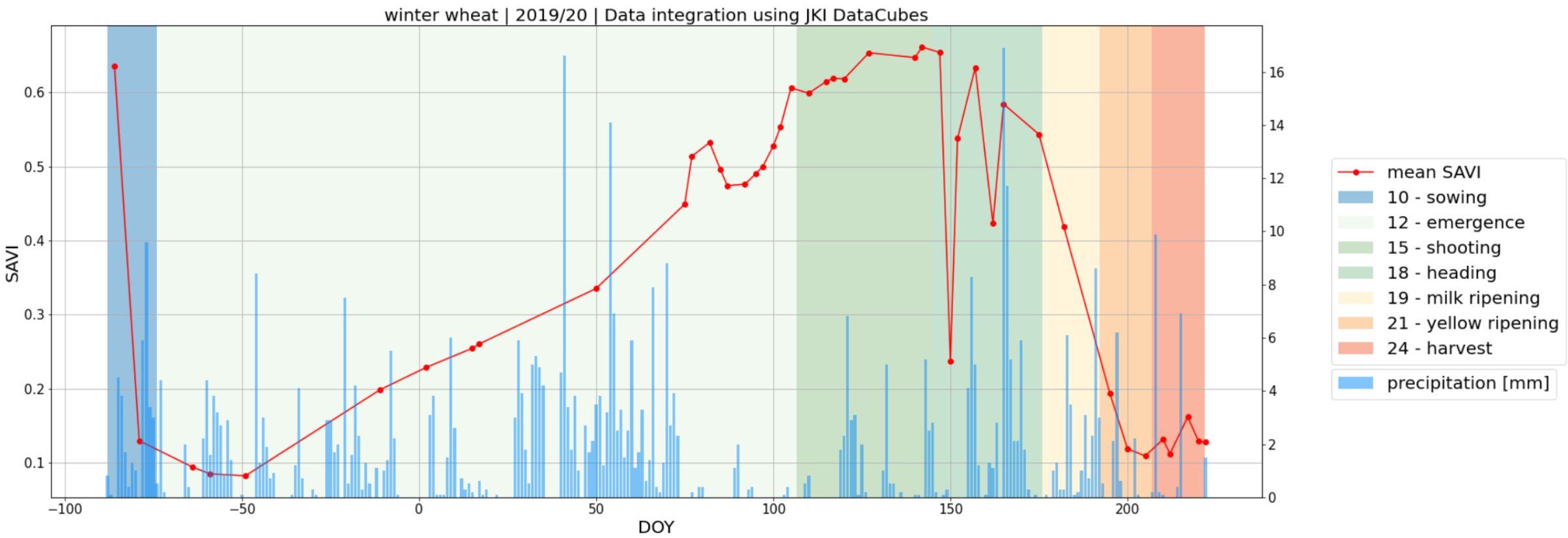
Soil factor



	ROT	NIR	SAVI
vital	0,02	0,45	0,66
gestresst	0,08	0,4	0,49

SAVI: Huete, A.R (August 1988). "A soil-adjusted vegetation index (SAVI)". Remote Sensing of Environment. 25 (3): 295–309. doi:10.1016/0034-4257(88)90106-X

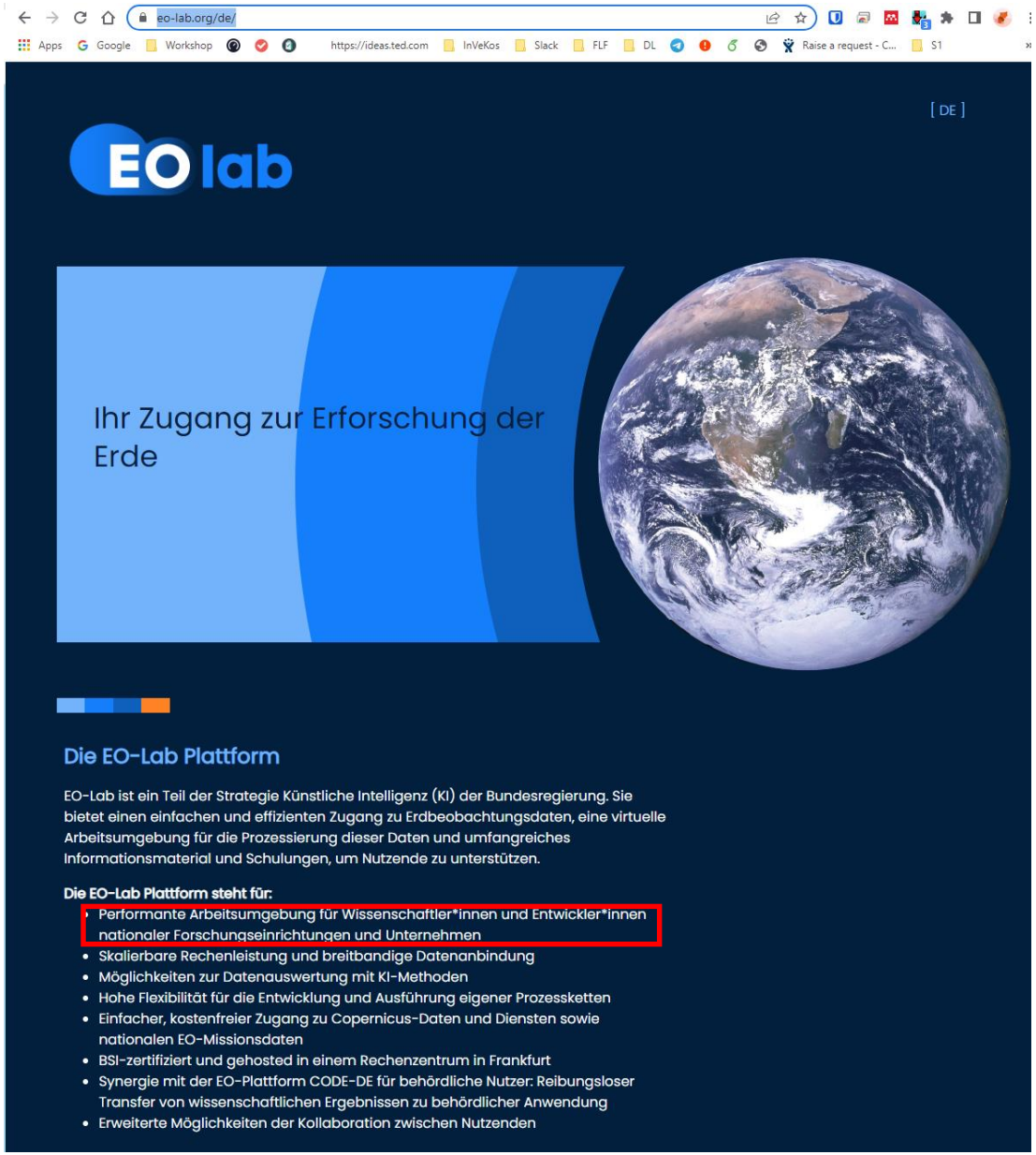
Demo Data Cubes



DEMO

Data Cubes

https://github.com/JKI-GDM/DemoPhaseWCS



eo-lab.org

KONTINGENTE können auch für Masterarbeiten erteilt werden!

Masterarbeiten am JKI?

Vergleich verschiedener Verfahren zur Wolken- & Wolkenschattenmaskierung auf Basis von Sentinel2

Spektrale Zonierung Deutschlands fokussiert auf landwirtschaftliche Flächen

Klassifikation und Spektralanalysen von Zwischenfrüchten mithilfe von Sentinel-1 und/oder Sentinel-2-Daten

Generalisierung phänologischer Phasen der Hauptfeldfrüchte Deutschlands

Meldet euch: florian.beyer@julius-kuehn.de

www.julius-kuehn.de

Danke für die Aufmerksamkeit

Kontakt: florian.beyer@julius-kuehn.de

esa.int/ESA_Multimedia/Videos/2016/08/Sentinel-2_global_coverage