



Earth Observation data
access and processing in the cloud
– new approach and info as a service



Dr Maciej Krzyżanowski
CEO, CloudFerro



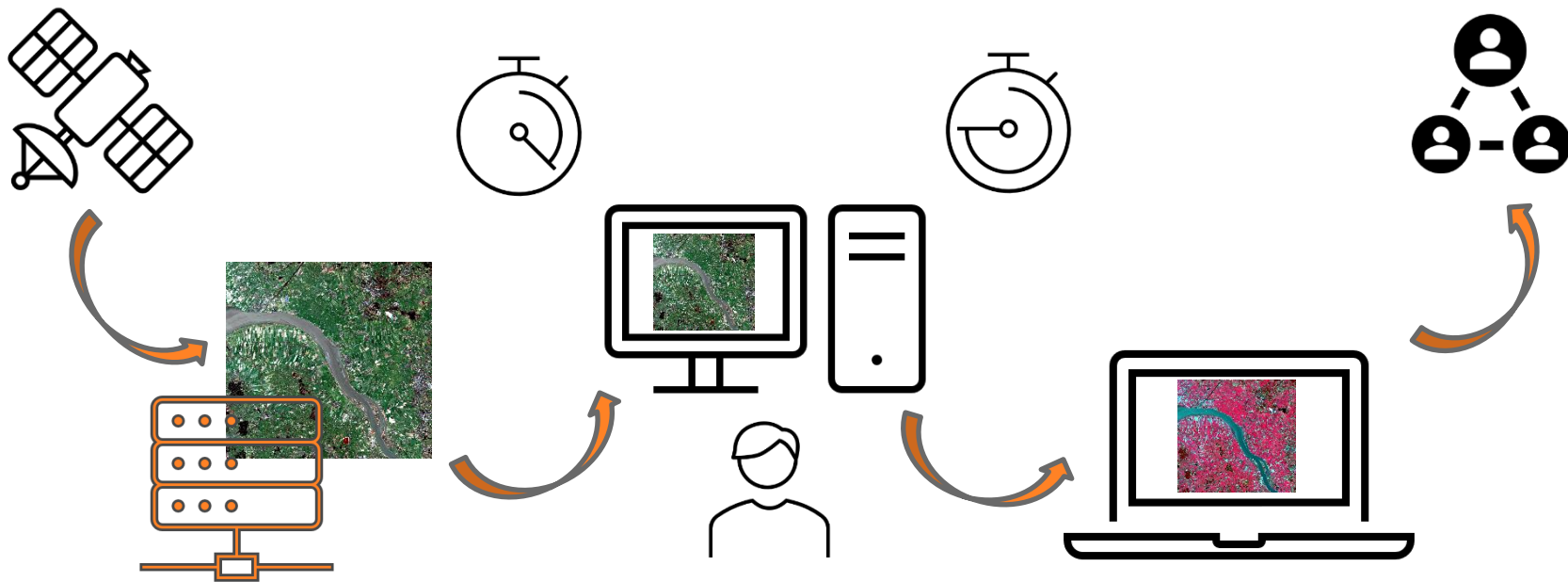
Agenda

1. EO data access and processing – **current challenges**
2. CREODIAS – **the new approach** for EO data and processing
3. Example **Use cases** of projects developed on **CREODIAS**

01

Earth Observation data access
and processing - **current challenges**

From data access to data sharing



EO data access and processing – challenges

- **data access** – where to download the data
- **data access** – where to obtain the data from different sources
- **data access** – how to ensure convenient data access for different use cases
- **huge data sizes** – how to acquire, store efficiently, ensure quality, keep costs low
- **data processing** – how to enable fast mass processing at scale
- **data processing** – how to facilitate repeatable processing to extract info

Why do we need EO Data access and processing platforms

Lots of data to be processed

Multiple, dispersed users

Users spend 80% of their time downloading and preparing data for processing

Massive processings, repeatable tasks – done quasi manually



Use a common data repository & cloud platform

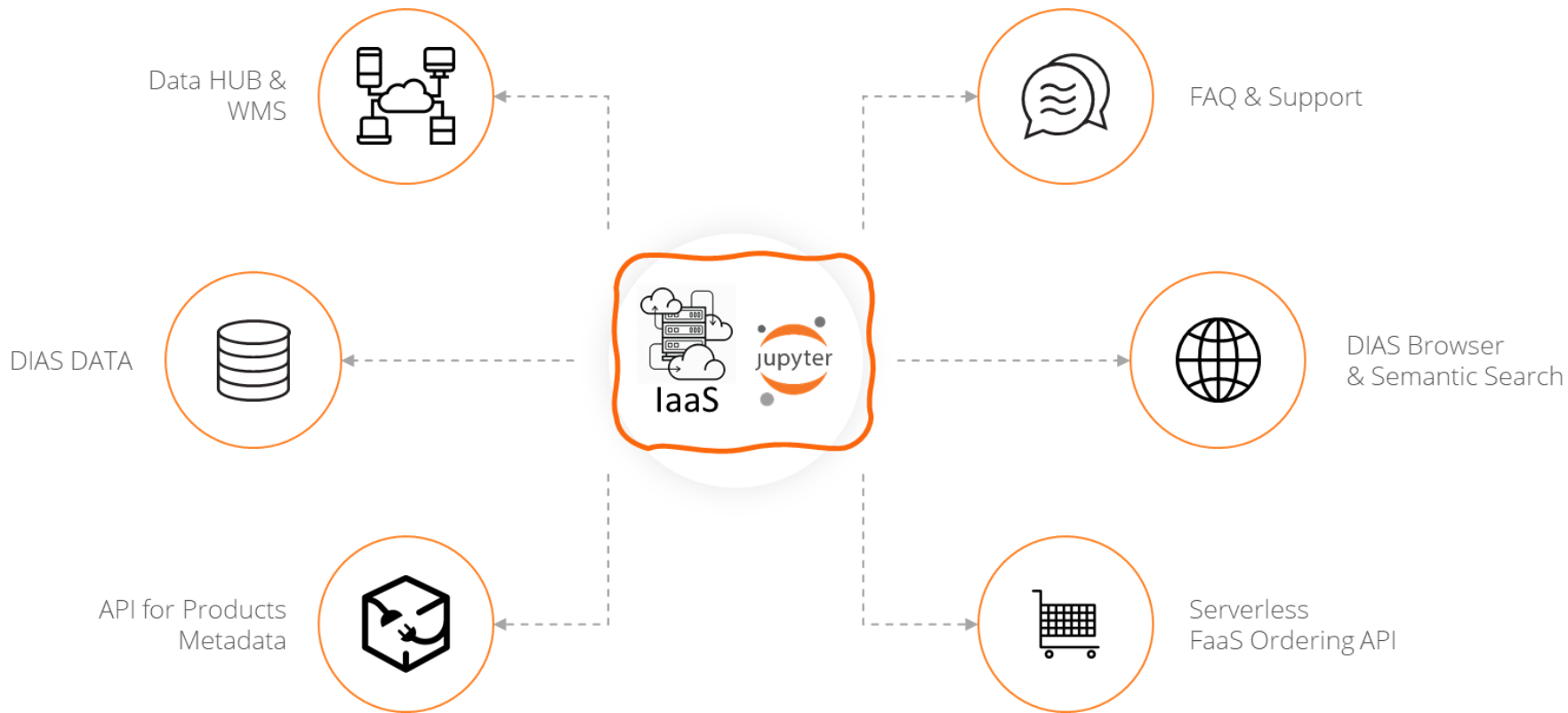
Bring processing to the data

Infrastructure as a code – a way to introduce automation and deal with massive and repetitive processings

02

CREODIAS platform - **the new approach**
to EO data access and processing

Data processing on a data platform - DIAS



Computing Resources

Virtual Machines and Storage
(all flavors available including nVidia GPU-based)

On-demand Processing (PGaaS)

JupyterHub

The screenshot displays two web interfaces. The top interface is the CloudFerre console, showing a list of instances under the 'Instances' tab. The bottom interface is the JupyterHub web application, showing a file explorer for the 'jupyter' environment.

CloudFerre Console - Instances

Instance Name	Image Name	IP Address	Flavor
aoe_test5	CentOS 7 Jul 2016-08-02-CLONE	private_network_00915 192.168.0.13 Floating IPs: 185.178.84.30 eodata 10.111.0.76	ec2.large

JupyterHub Interface

Files Running Clusters Nbextensions

Select items to perform actions on them.

Name	Last Modified	File size
..	seconds ago	
data	5 months ago	
docs	5 months ago	
libs	5 months ago	
src	5 months ago	
tests	5 months ago	
N1_Introduction.ipynb	Running 5 months ago	4.49 kB
N2_ObjectStorage.ipynb	5 months ago	323 kB
N3_SimpleDatabaseQueries.ipynb	5 months ago	93.4 kB
N4_QueryAndTransferCARDMetadata.ipynb	5 months ago	28.9 kB
N5_PlottingCARDSignatures.ipynb	5 months ago	164 kB
N6_PolygonAnalysis.ipynb	5 months ago	8.92 kB
N7_InteractiveMaps.ipynb	5 months ago	24.5 kB
N8_WorkingWithSentinel2.ipynb	5 months ago	946 kB
LICENSE	5 months ago	0 B
README.md	5 months ago	6.09 kB

CREODIAS Cloud Services (IaaS)

Cloud Services (IaaS)

FULLY
INTEGRATED
WITH EO
REPOSITORY
AND
CATALOGUE

BASED ON
INDUSTRY
STANDARD -
OPENSTACK

ACCESSIBLE
VIA API
AND GUI

SEPARATE
ADMINISTRATIVE
DOMAINS FOR
DIFFERENT
CUSTOMERS

PROCESSING :

- ✓ Virtual Machines
- ✓ Dedicated Server VM-s
- ✓ VM-s with GPU
- ✓ Containers
- ✓ FaaS

STORAGE :

- ✓ Block & Object
- ✓ HDD/SSD/NVMe
- ✓ Fully redundant
- ✓ Snapshots
- ✓ Backups

OTHER SERVICES :

- ✓ Virtual Networks & Routers
- ✓ Firewalls
- ✓ Floating IP
- ✓ DNS
- ✓ VPN
- ✓ Load Balancer

CREODIAS – over 23 PB of available EO data

1. Sentinel Missions
2. Copernicus Services
3. VHR Imagery (Jilin-1 / KazEOSat / KOMPSAT, Airbus.
4. ESA/Landsat Mission
5. MODIS Terra/Aqua
6. Envisat
7. SMOS
8. Jason-3
9. DEM
10. S2GLC



REPOSITORY

- High capacity object storage for EO Data (currently over 23PB)
- Data stored in its original, uncompressed form – direct access to individual files
- Accessible over NFS and S3 protocols
- Direct file visibility in user VMs – no need to download
- Dynamic generation of OGC WMS/WMTS/WCS tiles (integrated Sentinel Hub)
- Storage of public and private data
- 1 PB cache area for data ordered or generated by users
- Possible dissemination to remote users

EO DATA ACQUISITION

- Over 20 external EO data sources
- HTTP(S), (S)FTP, S3 and physical media transfer
- Optimized acquisition algorithm (parallelization, selection of fastest source)
- Fault tolerance and quality control
- Full data or metadata-only
- On demand acquisition (ordering)
- Subscription API for new products
- Customer-provided collections (restricted use)

AUTOMATION

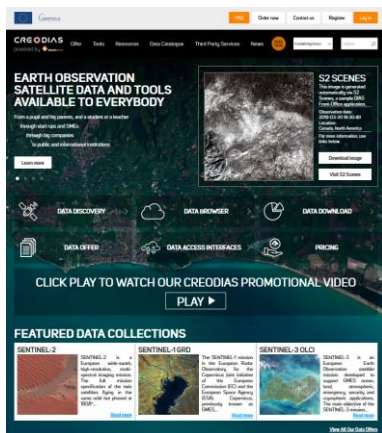
- All IaaS and PaaS functions available in m2m format via APIs
- Infrastructure as a Code – building complicated processing environments via programming and scripts
- Possible time scheduling of actions
- Chaining of applications – results from one can be inputs into another
- Flexibility

SOME APPLICATIONS

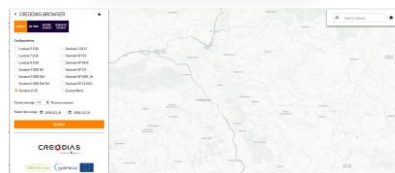
- Automated, repetitive processing
- E.g. Land cover classification every quarter or year or classification of crops
- Easy change detection
- Automated search for anomalies
- Additional analysis of selected anomalies
- AI training
- And many more.....

CREODIAS – ready to use user tools in the cloud

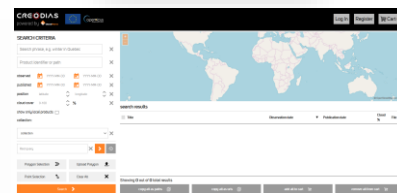
CREODIAS
Portal



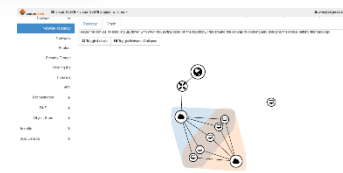
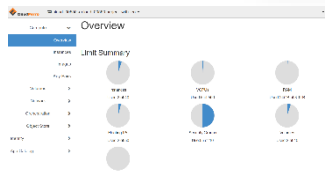
EO Browser



EO Finder



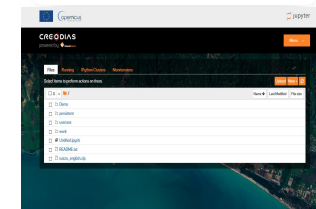
Cloud Dashboard



Third party
Applications



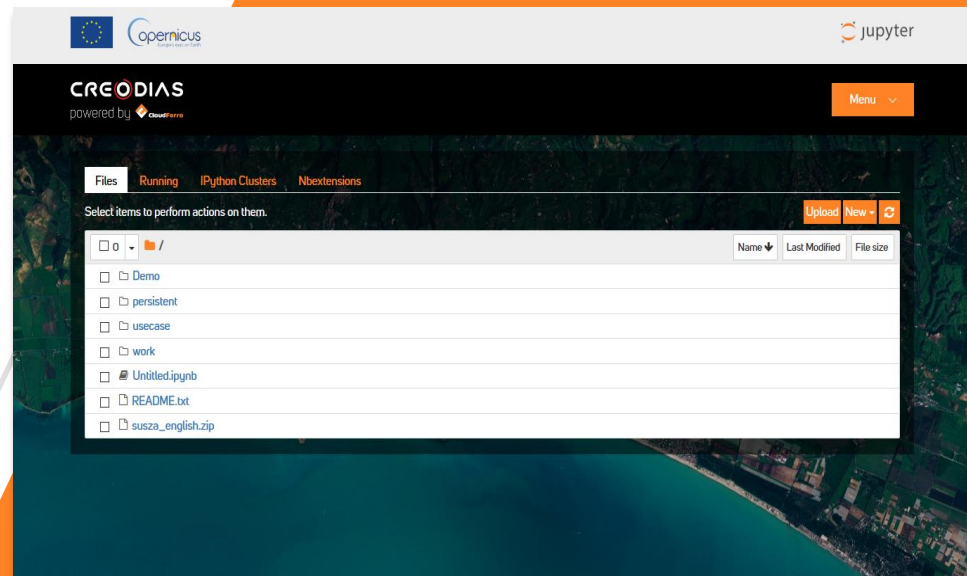
Jupyter Notebook



CREODIAS Jupyter Notebooks

Jupyter Notebook

- ✓ Free prototyping tool
- ✓ Ready-to-go
- ✓ Direct access to Earth Observation data repository
- ✓ Interactive web-based processing environment



03

Example Use Cases

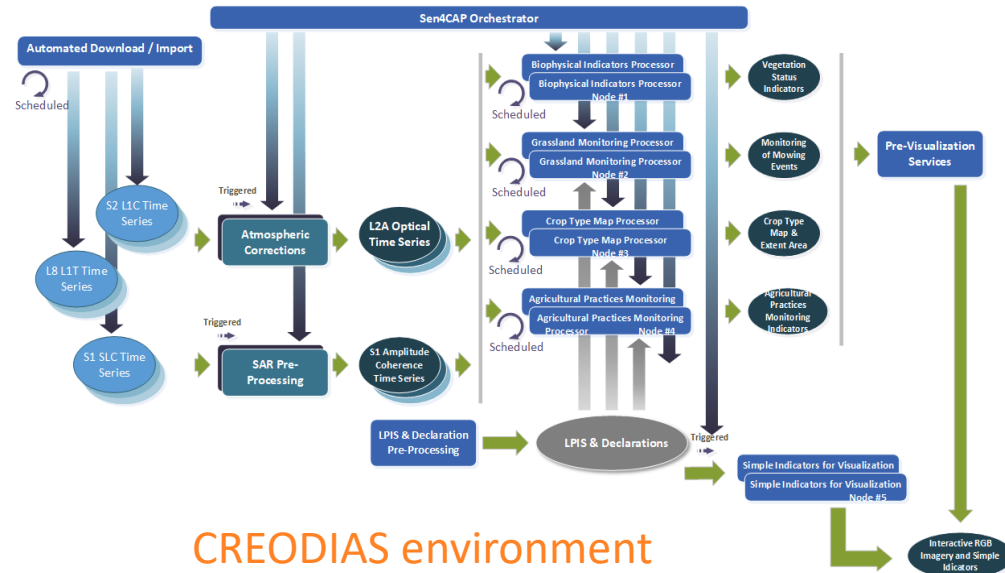
- projects developed on **CREODIAS**

Sen4CAP software

Ready-to-use monitoring solution for modern agriculture.

Based on EO data **Sen4CAP** allows to generate products supporting agricultural analyses:

- ▶ Cultivated **crop type** map
- ▶ **Grassland mowing** product
- ▶ Vegetation **status** indicator
- ▶ Agricultural practices monitoring products



S2GLC – Sentinel-2 Global Land Cover classification



LAND COVER MAP

S2GLC Extension – Europe

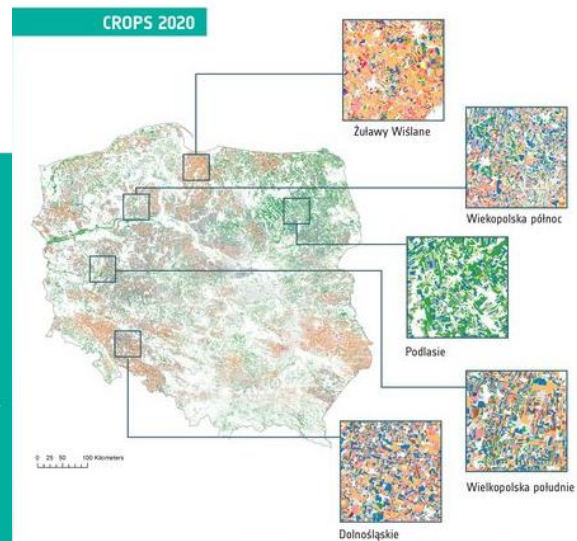
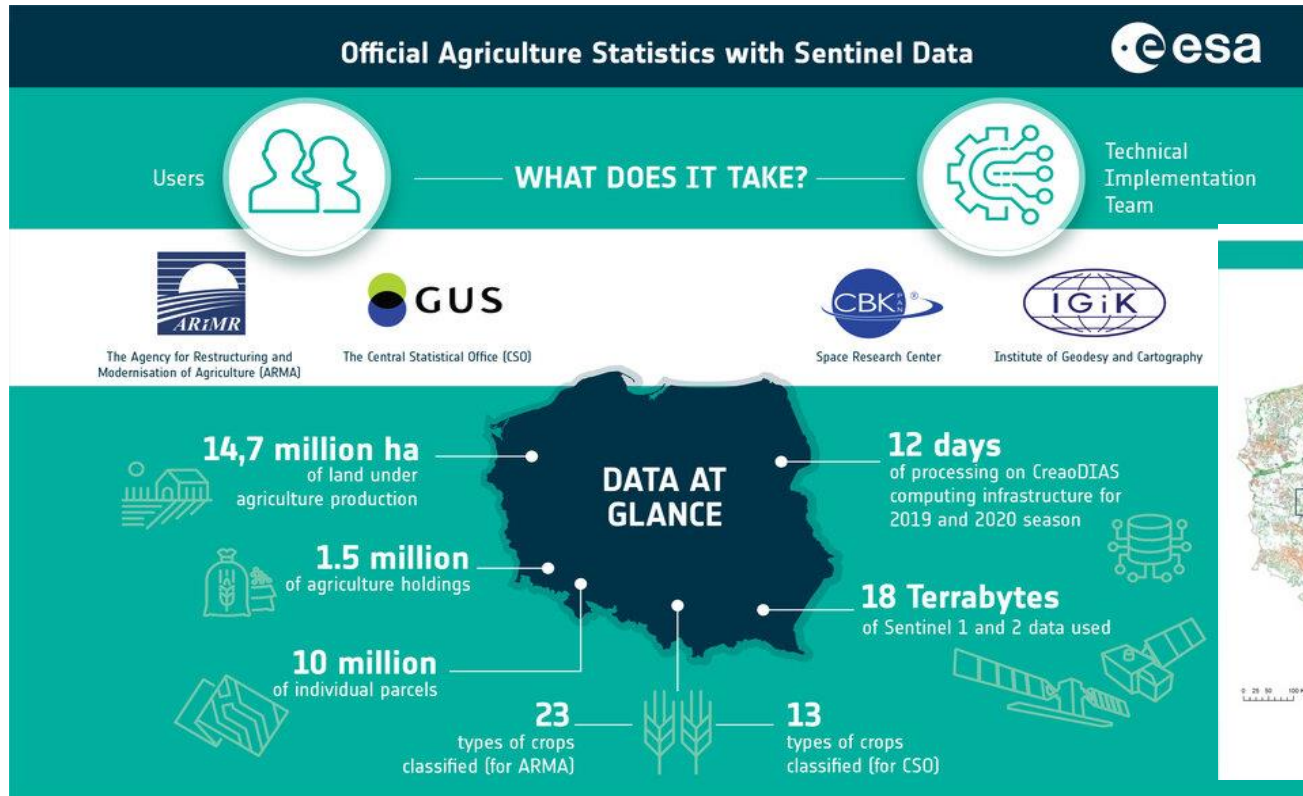
Classification of Sentinel-2 data from year 2017

LEGEND

- | | |
|---------------------------------------|---------------------------------|
| 5-2 tile | Herbaceous vegetation |
| Clouds | Moors and Heathland |
| Artificial surfaces and constructions | Sclerophyllous vegetation |
| Cultivated areas | Marshes |
| Vineyards | Peatbogs |
| Broadleaf tree cover | Natural material surfaces |
| Coniferous tree cover | Permanent snow covered surfaces |
| | Water bodies |



EOStat: Services for Earth Observation-based statistical information for agriculture



Thank you for your attention

Maciej Krzyżanowski
mkrzyzanowski@cloudferro.com

For more information,
please visit: www.cloudferro.com

Link up with us on:

 <https://twitter.com/CouldFerro>

 <https://www.linkedin.com/company/clfr/>

 <https://www.facebook.com/cloudferro/>

