

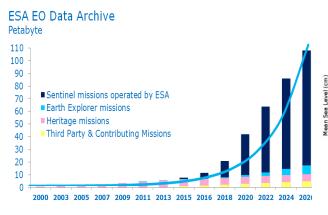
# The EO platform ecosystem addressing the Green Deal: The ESA Green Transition Information Factories

08 December 2022 – EuroGEO Workshop 2022, Athens, GR Patrick Griffiths – ESA EOP - Science, Applications & Climate Department



## **Big Data Challenges in EO**





B

Mutil-missions MSL

Topex/Poseldon

Jason-1

Jeson-2

Envisat

ERS2

GFO

O CNES,ESA,LEGOS,CLS

2

1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014

**Data continuity** 





**Data volume** 

The state of the s

**Data sharing** 

March 2014

With SMOS

Nov. 2014

Differences of about 20cm

Man NCC

SMON 2014

Jan NCC

Jan

**Data quality** 

**Innovation** 

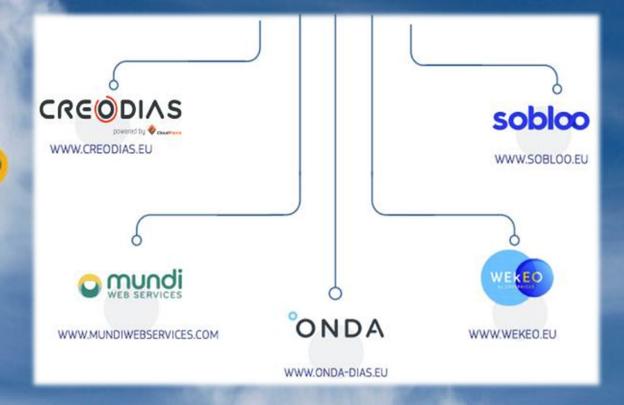
**Timeliness** 

Mission synergies

**Diversity** 

## "Move the algorithm/user to the data"

## PARADIGM AHEAD

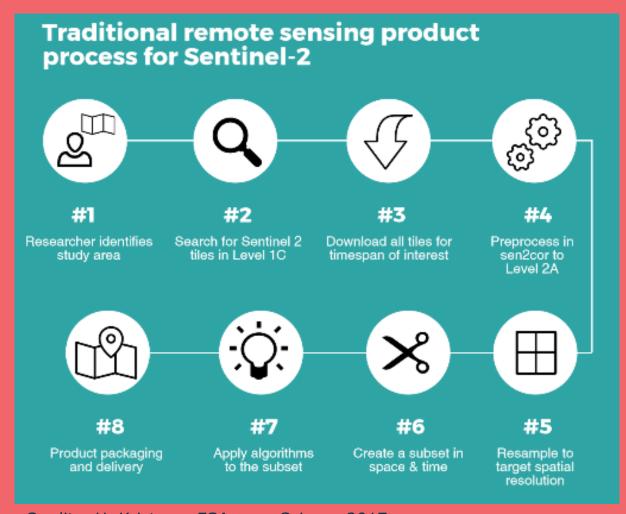


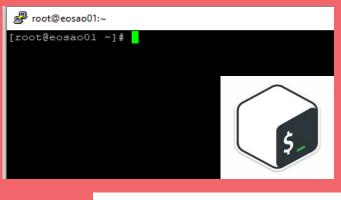
...maybe it's not that simple:

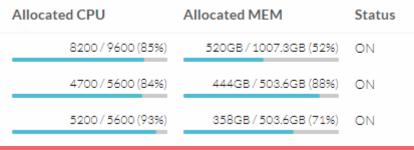
http://storyofchange.net

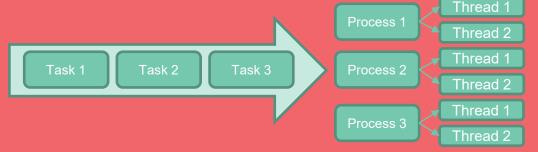
## The EO Data Management Burden...











Credits: H. Kristen - ESA open Science 2017

## EO Platforms: innovation for science & applications



## Simplification & democratisation of EO:

- Streaming EO data, pixel-level access
- Dynamically allocating compute resources
- Intuitive front end syntax, operator chaining
- Reusable mass processing pipelines

#### Scientific advances:

- Facilitating reproducibility
- Paper & code
- Transparency
- Innovation

## Collaborating and sharing:

- Sharing user of code, prevent reinventing the wheel
- Provide optimised processes as building blocks
- Expose mature algorithms as on-demand services

### Interactive development:

- Ad-hoc visualisation of intermediate results
- Hosted Jupyter notebooks
- Deferred (lazy) evaluation
- Executable tutorials

Slide 5



















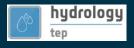


















































































Earth

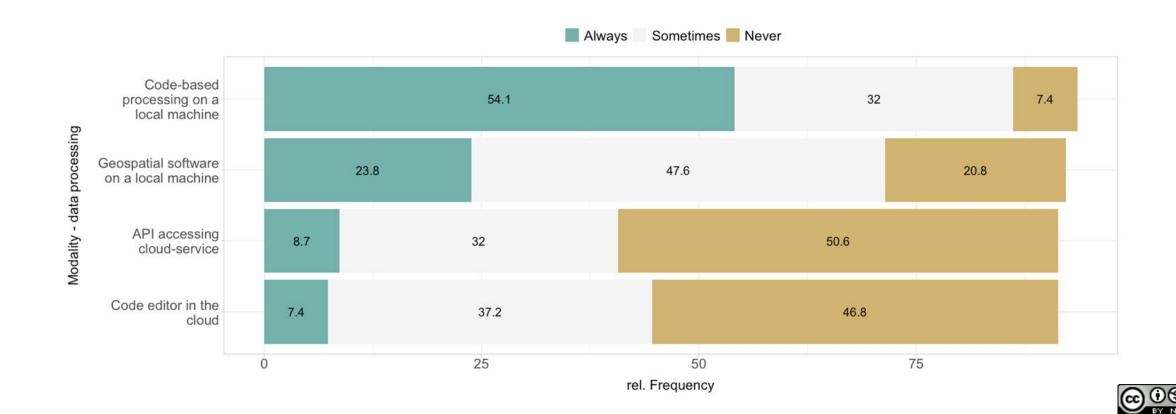
aws

#### J. Wagemann et al, 2021: "Users of open Big Earth data", https://doi.org/10.1016/j.cageo.2021.104916

Code-based processing on a local machine (86%) is prevailing data handling modality

70% further indicated to use always or sometimes a locally installed geospatial software additionally

Less than half handle and process data with a cloud-based service



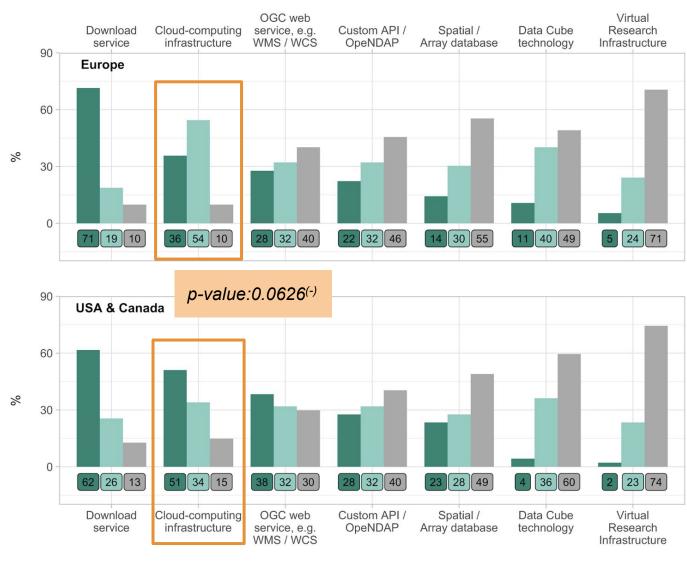


Use pattern between Europe and USA & Canada is significantly different at the 0.1 significance level

Current uptake of cloud services seems to be higher in the USA & Canada

In Europe, fewer are already using a cloud service, but more than half is interested in using it in the future

Data system	Chi-square / Degrees of Freedom (DF)	<i>p</i> -value
Download service	1.459/2	0.4821
Cloud-computing infrastructure	5.5412/2	0.0626(-)
OGC web service	2.1754/2	0.337
Custom API / OpenDAP	0.59166/2	0.7439
Virtual Research Infrastructure	0.86255/2	0.6497
Data cube technology	2.3999/2	0.3012
Spatial / Array database	1.9579/2	0.3757



Interest in future use

No interest

Currently used

## EO Platforms - Capability gaps in Europe



The EO cloud processing/analytical landscape in Europe is not matching European leadership in EO observational capabilities:

- Fragmentation, redundancy and lacking coordination among **EO platform providers** (and funding bodies)
- Prevalence of **old Virtual Machine model** (rather than dynamic resource allocation and scaling);
- File-based storage and data access, rather than pixel-level flexibility and cloud native data structures;
- Unappealing business models, lacking long term perspectives, project centric thinking

## openEO Platform



#### openEO Platform:

- A new cloud-based, federated and open-source EO analytics environment
- JupyterLab and web-editor for interactive development
- Intuitive front-end libraries in Python, JavaScript and R 🔁 🥦



- Demonstrates European excellence building on achievements of the H2020 openEO API
- openEO platform aims for optimal conformity with requirements of real-world EO users



Kick-off September 2020

Operational service launched at 2021 ESA Phi Week

User Consultation and classroom trainings at ESA LPS 2022



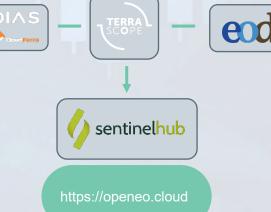




Community & Federation



Providing transparency



OpenE0

**Platform** 



















## openEO platform - Process graphs



A openEO process graph is a chain of specific processes that organize and automate the execution of one or more processes. They are defined in JSON and are therefore

programming language agnostic.

```
load collection #dc
                                                     Enhanced Vegetation Index (EVI)
id: "Sentinel-2"
                                                                                                            array_element #blue
spatial extent:
                                                                                  array element #red
                                                                                                            data: data from #evi
       west: 16.1
                                                                                  data: data from #evi
       east: 16.6
                                                                                                            index: 2
                                                       array_element #nir
                                                                                  index: 1
       north: 48.6
                                                       data: data from #evi
       south: 47.2
                                                       index: 0
temporal extent: ["2018-01-01", "2018-02-01"]
                                                                                         product #p1
                                                                                                             product #p2
bands: ["nir", "red", "blue"]
                                                                                        data: [6, #red]
                                                                                                             data: [-7.5, #blue]
                                                       subtract #sub
                                                                                     sum #sum
                                                       data: [#nir, #red]
                                                                                    data: [1, #nir, #p1, #p2]
            reduce #evi
            data: #dc
                                                         product #p3
                                                                                   divide #div
            dimension: "spectral"
                                                        data: [2,5 #div]
                                                                                   data: [#sub, #sum]
            reducer: callback >
            reduce #mintime
           data: #evi
                                                     data: data from #mintime
           dimension: "temporal"
           reducer: callback >
               save result
               data: #mintime
               format "GTiff"
```

```
"dc": {
  "process_id": "load_collection",
 "process_description": "Loading the data; The order of the specified bands is
 "arguments": {
   "id": "Sentinel-2",
    "spatial_extent": {
      "west": 16.1,
     "east": 16.6,
      "north": 48.6,
      "south": 47.2
    "temporal_extent": ["2018-01-01", "2018-02-01"],
    "bands": ["B08", "B04", "B02"]
"evi": {
 "process_id": "reduce",
 "process_description": "Compute the EVI. Formula: 2.5 * (NIR - RED) / (1 + NIR
 "arguments": {
   "data": {"from_node": "dc"},
    "dimension": "spectral",
   "reducer": {
      "callback": {
        "nir": {
          "process_id": "array_element",
          "arguments": {
            "data": {"from_argument": "data"}.
            "index": 0
```

## openEO platform - ARD processing



#### openEO processes provide optimised, high-level operators as building blocks

```
import openeo
connection = openeo.connect("openeo.cloud").authenticate oidc()
S2L1C = openeo cloud.load collection(collection = 'SENTINEL2 L1C',
                                     spatial_extent = {'west':bbox[0],'east':bbox[2],'south':bbox[1],'north':bbox[3]},
                                     bands=["B02","B03","B04","B08","CLM",'B09','B8A','B11'],
                                     temporal extent=temporal extent)
S2ARD = ard_surface_reflectance(s2_l1c,
                               'FORCE',
                               'FMASK'.
                               elevation model = 'COPERNICUS 30',
                               atmospheric correction options = {'DO BRDF=TRUE','DO TOPO=True'},
                               cloud detection options = {'CLD PROB=0.15','CLD DIL=6','SHD DIL=6'})
```

## openEO platform - resources



- Platform portal: <a href="https://openeo.cloud/">https://openeo.cloud/</a>
- Web editor: <a href="https://editor.openeo.cloud/">https://editor.openeo.cloud/</a>
- Jupyter lab: <a href="https://lab.openeo.cloud/">https://lab.openeo.cloud/</a>
- user forum: <a href="http://forums.openeo.cloud/">http://forums.openeo.cloud/</a>
- documentation: <a href="https://docs.openeo.cloud/">https://docs.openeo.cloud/</a>
- GitHub: <a href="https://github.com/orgs/openEOPlatform/repositories">https://github.com/orgs/openEOPlatform/repositories</a>
- docs section on virtual datacubes:
  <a href="https://openeo.org/documentation/1.0/datacubes.html">https://openeo.org/documentation/1.0/datacubes.html</a>
- openEO (h2020) cookbook:
  <a href="https://openeo.org/documentation/1.0/cookbook/#input-load-collection">https://openeo.org/documentation/1.0/cookbook/#input-load-collection</a>







- Curation for training data, ground truth, pre-trained models, benchmarks;
- Provide relevant tooling for all TDS related needs
- Integrate with and build on top of other EO platform assets (e.g. for feature engineering, large scale inference, etc.);
- Explore Quality Assurance, Community Incentivization, etc.











- Kicked off in September 2022 for an initial year contract
- Please get in touch if you are interested in this topic!





## EARTH OBSERVATION TRAINING DATA LAB (EOTDL)

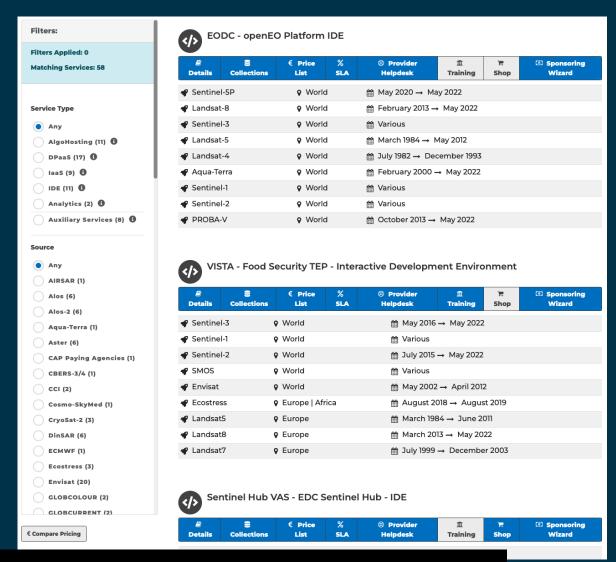
https://eotdl.vercel.app/

## **ESA Network of Resources**

## http://nor-discover.cloudeo.group/



- Allow users to discover suitable resource tier and platform tier services to fulfil their needs.
- Single reference point providing information on providers including pricelists, helpdesk contact etc
- Link to the providers for users who wish to make commercial purchases
- An embedded sponsorship form
  - No registration required
  - Clear maximum sponsoring limits per service and overall project limits shown
- Comparison of laaS services including factors like Research discounts applied to prices



More than 530 projects from 79 Countries approved for sponsoring

## **EO Platform Ecosystem – Lessons Learnt**



- Cloud based EO platforms can provide various abstraction levels that help scientist to cope and work more efficiently with the big EO data archives
- The European EO platform landscape suffers from redundancies and fragmentation:
  - Change of mind-set, away from project-centric thinking, towards interoperable building blocks, stop constantly "reinventing the wheel", rather reuse and build upon;
  - Key technology elements are emerging via community adaptation (e.g. STAC, openEO API);
  - Advanced federation concepts can help consolidation towards an interoperable ecosystem;
  - New Copernicus Data Access Service...
- There is large potential for EO and the EO platform ecosystem to support the implementation of the European Green Deal ambitions!

## EO in support of the European Green Deal



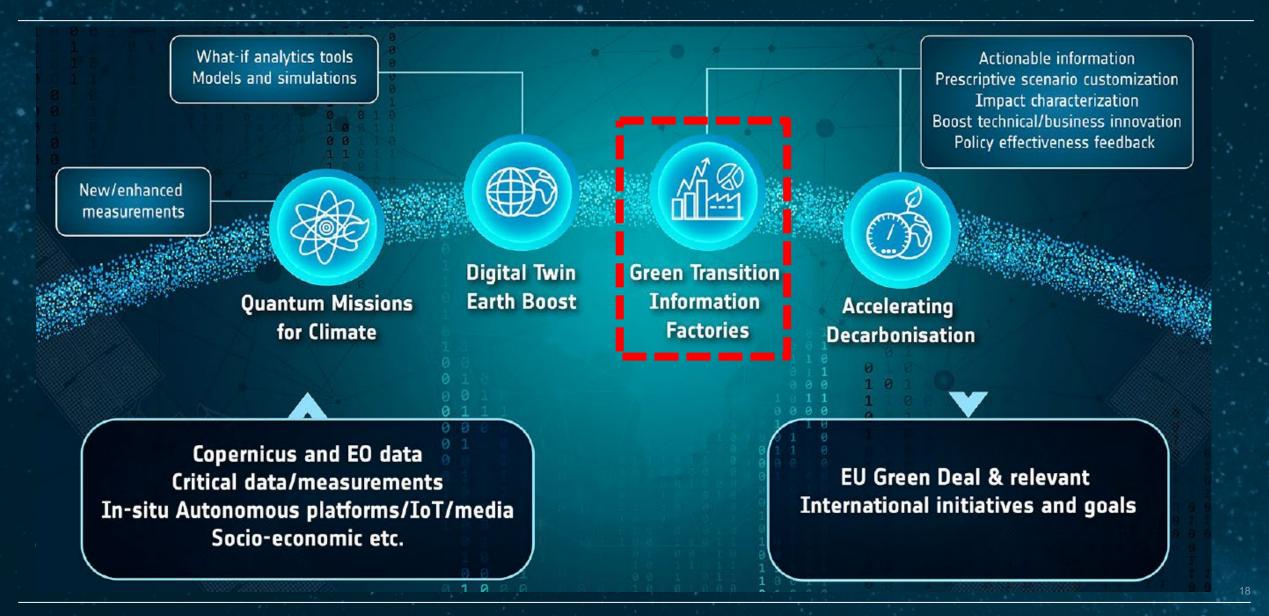
❖ The Green Deal is likely going to be the most impactful policy for the EO sector after the CAP...





## Space for Green Future (S4GF) Accelerator





## Green Transition Information Factory - Demonstrator for Austria



- Demonstrate added value of EO and digital platform technologies for addressing information needs in the context of the Green Transition and the European Green Deal;
- ❖ Coordinate with stakeholders and engage industry to match national priorities with capabilities;
- Provide innovative interactive tools and indicators within the Information Factory environment, supporting and enabling policy makers, analysts and citizens to engage in the Green Transition;

## **Capture User Requirements**



Address national priorities & information needs



Interactive GTIF explore tools for analysts, policy makers and citizens







## GTIF-AT: Involved companies and stakeholders



#### STAKEHOLDER CONSULTATION — IDENTIFY NATIONAL PRIORITIES





**ENVIRONMENT AGENCY AUSTRIA** 

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie

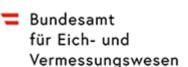














#### GTIF IMPLEMENTATION — RESPONDING WITH EO-BASED SOLUTION































