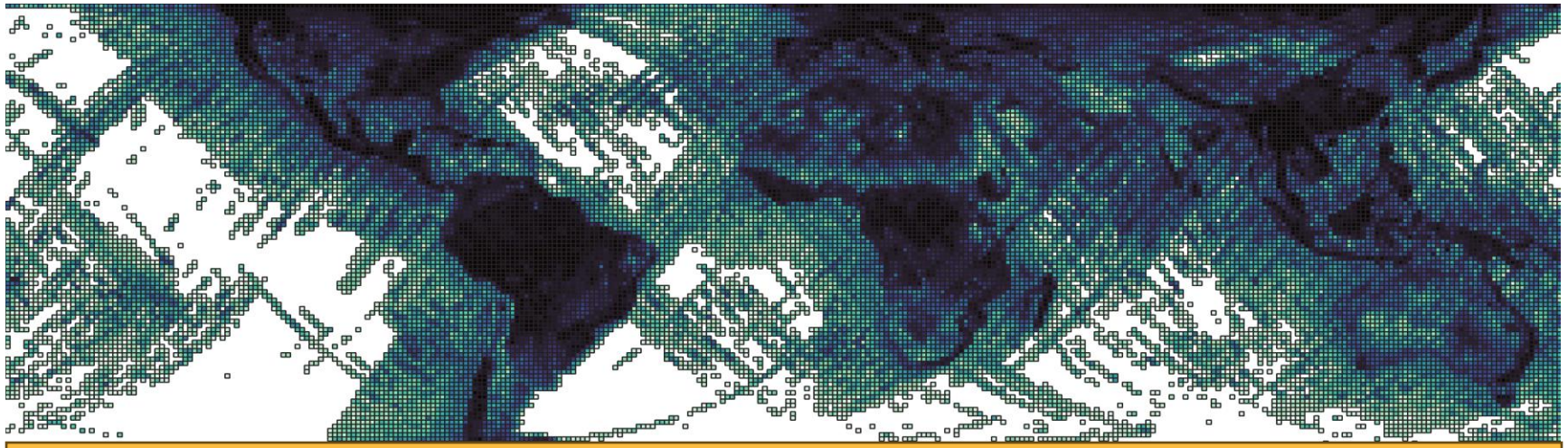
International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at



GEO-OPEN-HACK: An Initiative On Big Geospatial Data Processing With Open Computing Infrastructure And Open Tools

International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at

OPEN EARTH MONITOR

SPATIAL ECOLOGY

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101059548.

Milutin Milenković (milenkovic@iiasa.ac.at)

Novel Data Ecosystems for Sustainability (NODES), IIASA

Date: 2. October 2025

ESA BiDS Conference, Riga, Latvia

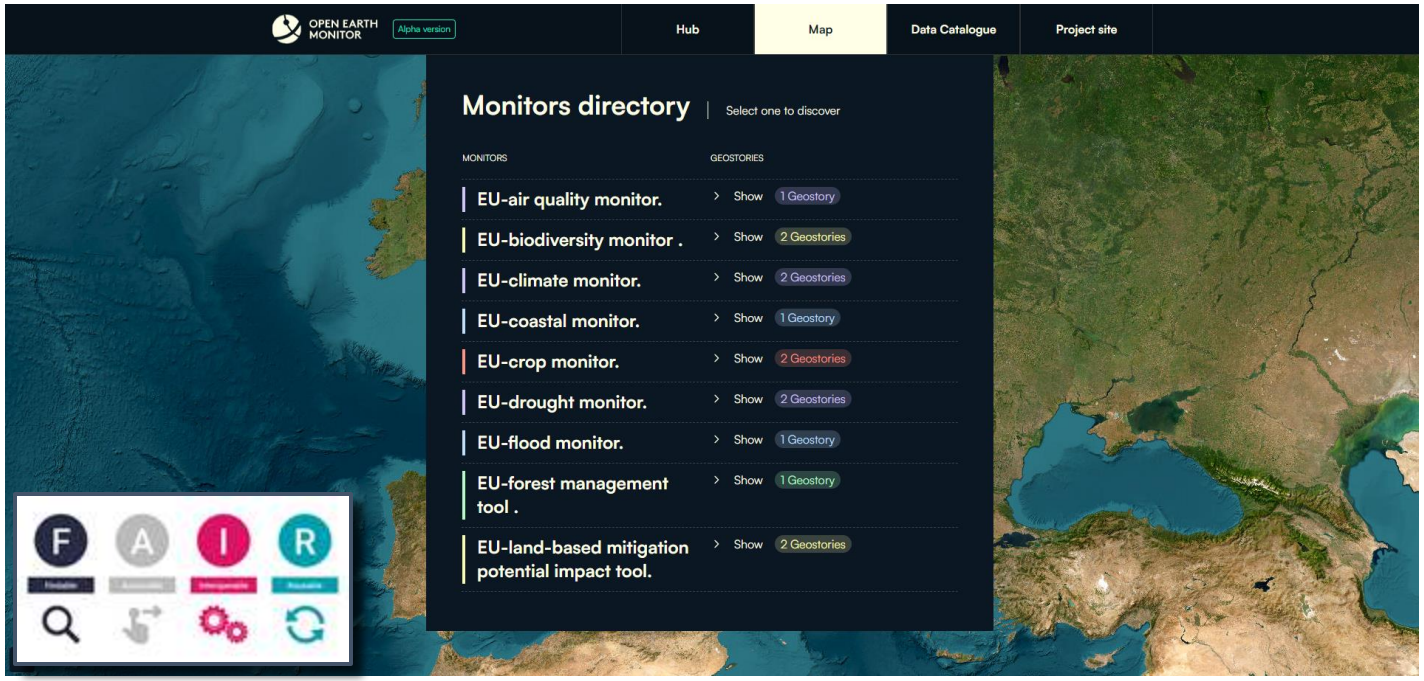
Open-Earth-Monitor Cyberinfrastructure (OEMC)

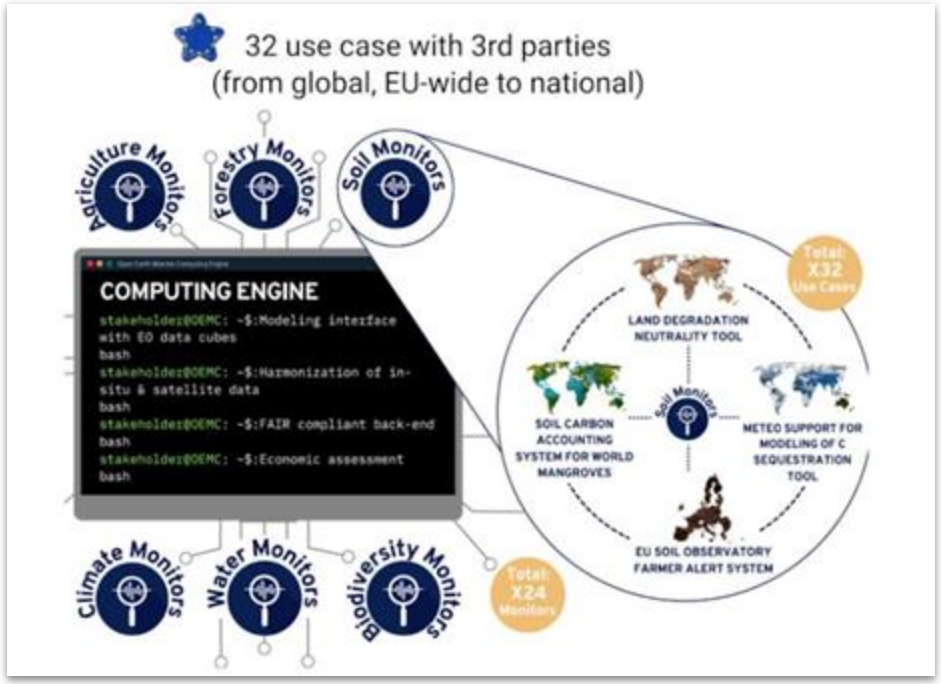
Consortium: 23 partner organizations
Funding: EUR 12 720 045,00
Period: June 2022 - July 2026
Call: Horizon CL6 2021 governance 01

FAIR Science: Findable, Accessible, Interoperable & Reproducible

1. **Reproducible research**: computational notebooks, knowledge hubs, …
2. **Open Source software**: AGPL, GPL, MIT, Apache, …
3. **Open Data license:** CC-BY, CC-BY-SA, ODbL, …
4. **Data & metadata repo's + catalogs:** STAC, Zenodo …

https://app.earthmonitor.org

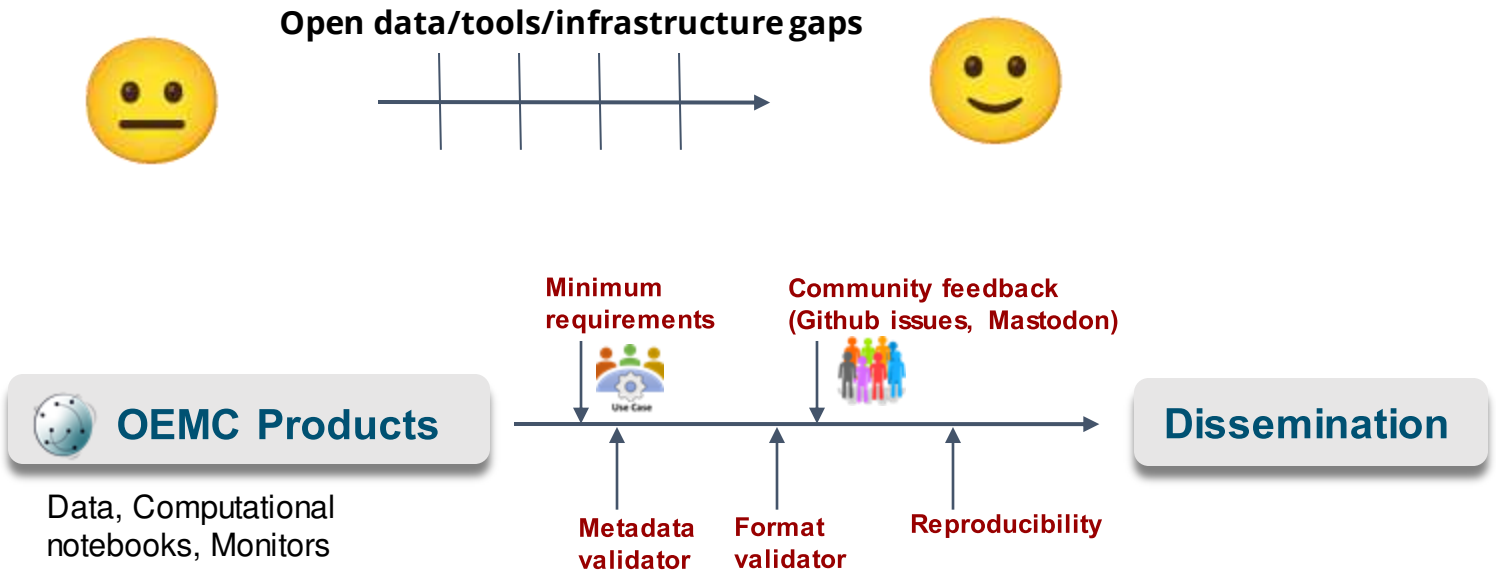


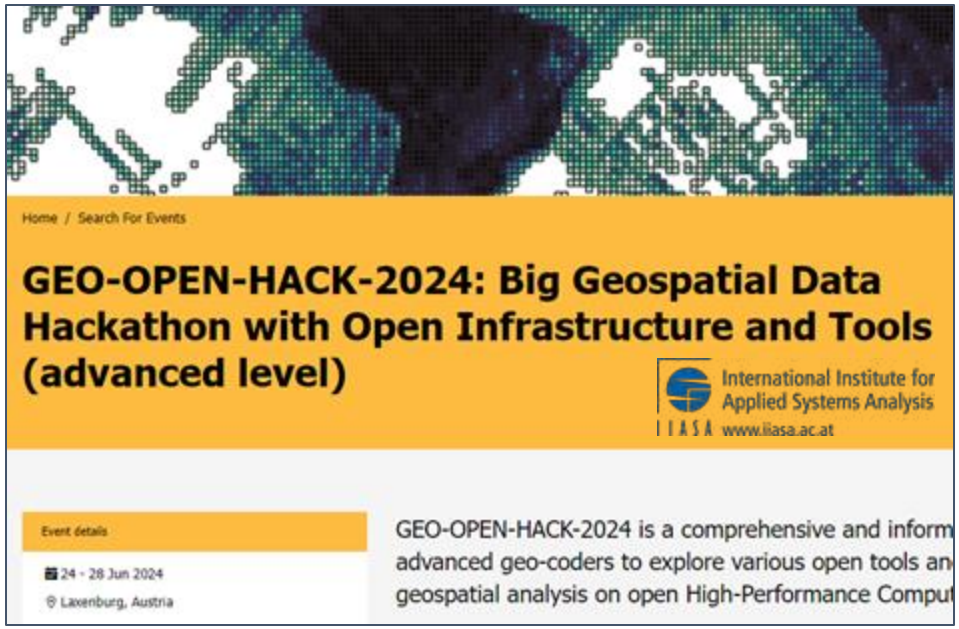


GEO-OPEN-HACK-2024: Aim & Goals

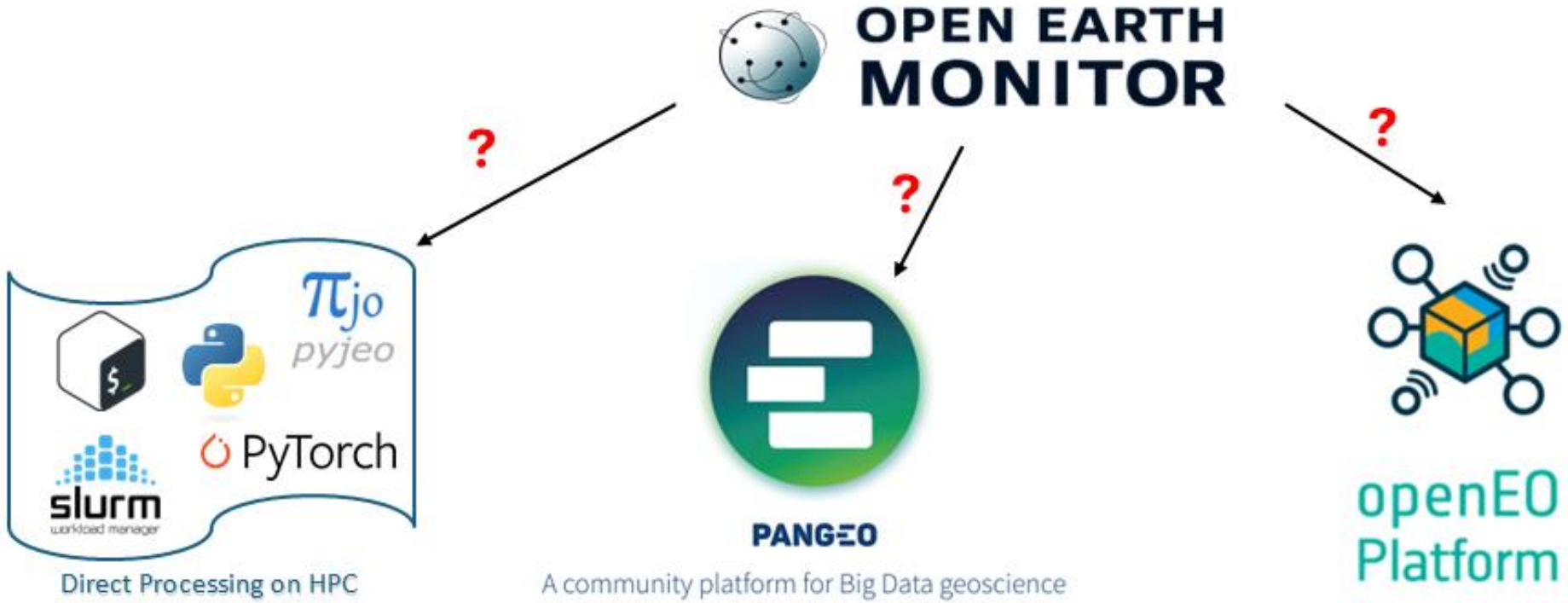
… is **a non-profit, give-and-take initiative** introduced under the OEMC project to:

- to lower the barrier and transfer knowledge to users dealing with big geospatial data analytics
- familiarize with and understand the differences among different big-data computing approaches
- present and get feedback on OEMC solutions/products
- identify current gaps in big data computing





Computing Approaches



Agenda GEO-OPEN-HACK 2024

OPEN EARTH MONITOR

Time	Sunday 23-Jun-24	Monday 24-Jun-24	Tuesday 25-Jun-24	Wednesday 26-Jun-24	Thursday 27-Jun-24	Friday 28-Jun-24
08.00-08.30					Canadian Pancake Breakfast IIASA Schlossrestaurant	
08.30-09.00		Get together IIASA WODAK meeting room	Get together IIASA WODAK meeting room	Get together IIASA WODAK meeting room		Get together IIASA WODAK meeting room
09.00-09.30		Keynote Raymond Oonk, SURF	Keynote Tom Hengl, OGH	Keynote Jan Verbesselt & Dainius Masiliunas	Keynote Edzer Pebesma, Uni Münster	Keynote Wolfgang Wagner, TU Wien & EODC
09.30-10.00		Giuseppe Amatulli Geo-Processing with HPC	Pieter Kempeneers Geo-python with HPC	Keynote Patrick Griffiths, ESA	Claus Michele & Valentina Premier openEO	Anne Fouilloux & Tina Odaka Pangeo
10.00-12.30				Antonio Fonseca ML with HPC		
12.30-13.30		Lunch break IIASA Gvishianhi meeting room	Lunch break IIASA Gvishianhi meeting room	Lunch break IIASA Gvishianhi meeting room	Lunch break IIASA Gvishianhi meeting room	Lunch break IIASA Gvishianhi meeting room
13.30-14.00		Big Data Story Giuseppe Amatulli	Big Data Story Milutin Milenković & Florian Hofhansl	Big Data Story Yu-Feng Ho & Johannes Heisig	Big Data Story Valentina Premier	Big Data Story Leandro Parente
14.00-17.30		Working with your data	Park Visit Laxenburg Park – START: 16.00h Working with your data	Working with your data	Working with your data	Open floor sessions OEMC and other open initiatives. OEMC data platform & feedback. Where to implement my big data problem?
18.00-21.00	Ice Breaker Event Rest. Stöckl im Park Prinz-Eugen-Strasse 25 1030 Vienna			Social Event IIASA Schlossrestaurant Herzog Albrecht-Strasse 1 2361 Laxenburg		

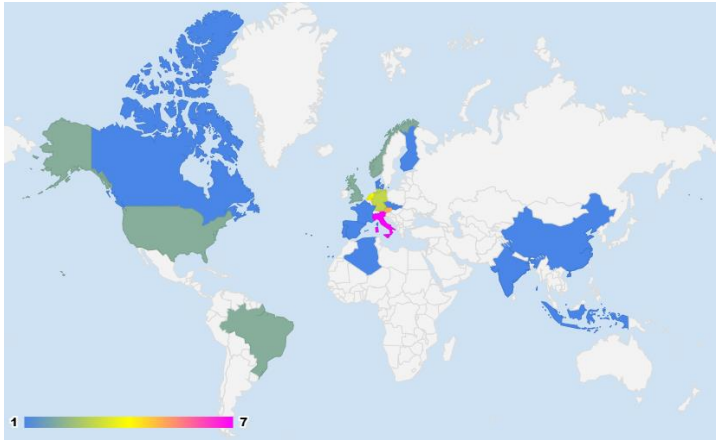
This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No.101059548

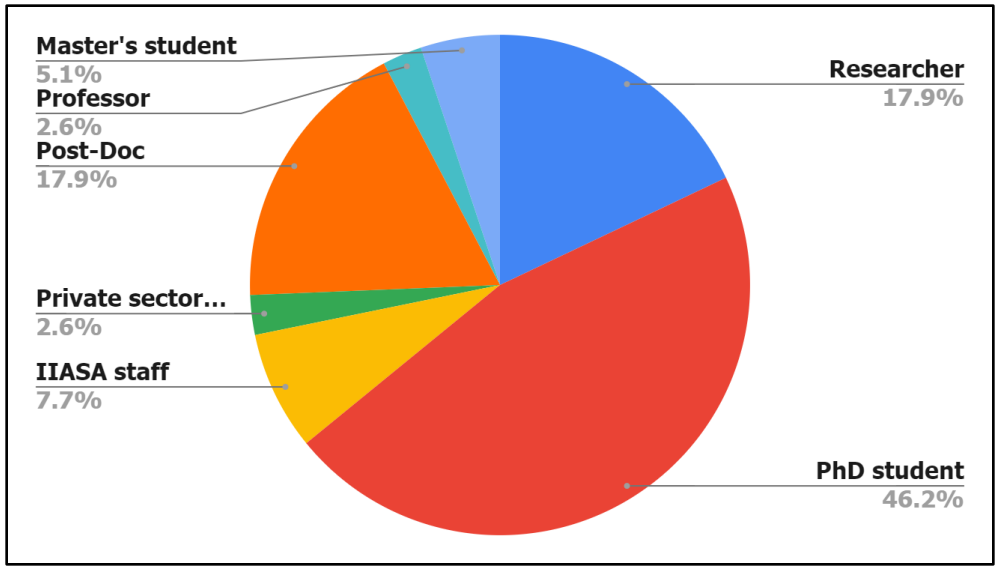
GEO-OPEN-HACK-2024 Participants

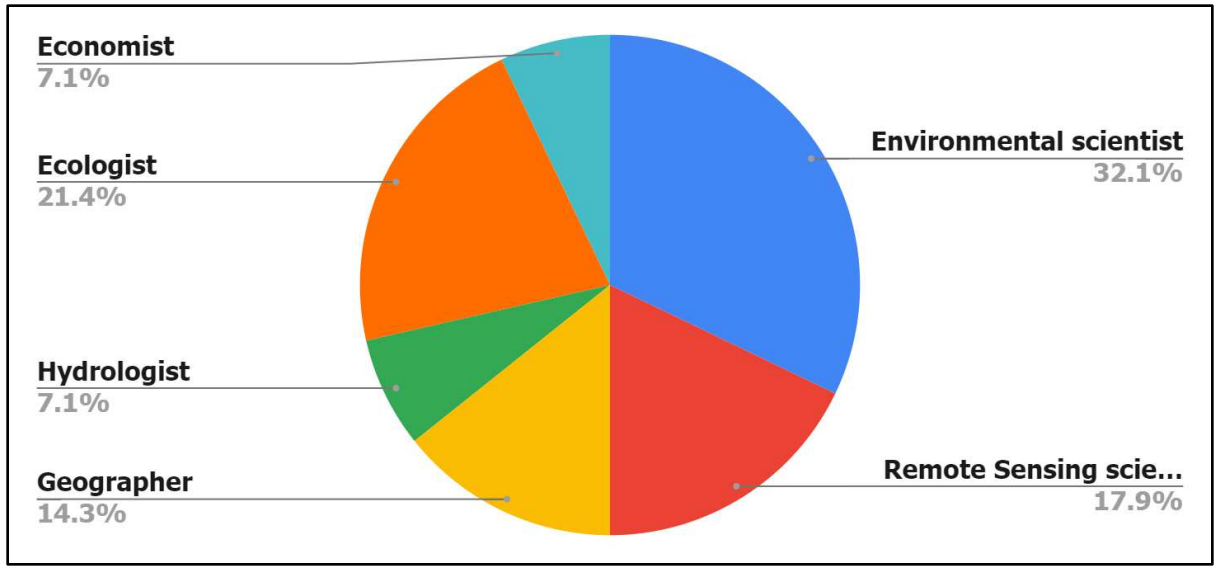
Participants.....41
- on-site.....20
- online.....21

Teachers.....11
Keynote Speakers.....7
Organization.....4

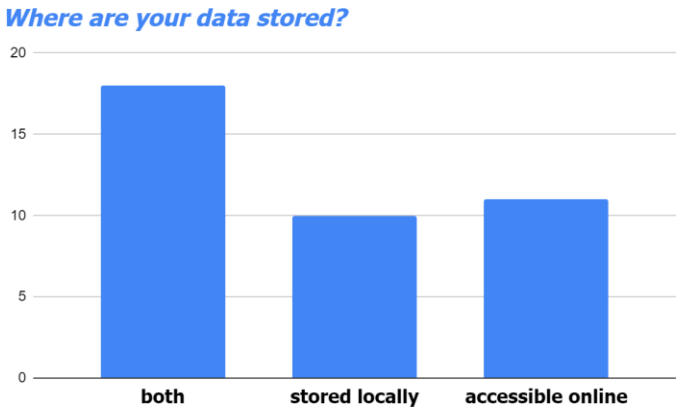
Total 63



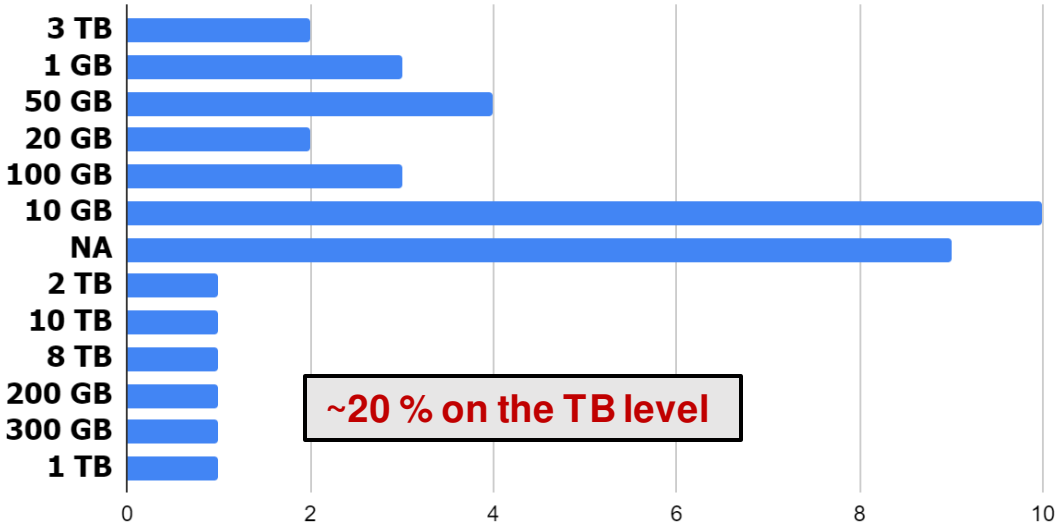




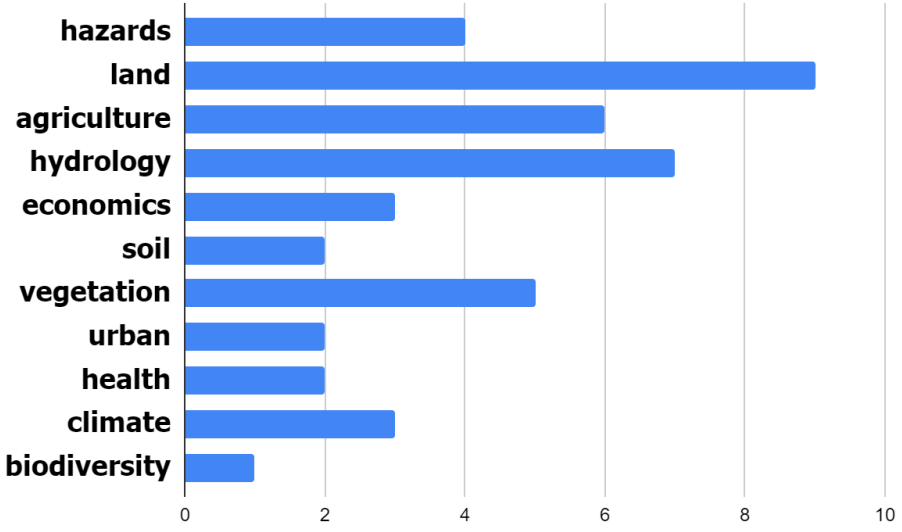
Big Data Project



If you have some data locally stored, please specify what is the total size of this dataset?

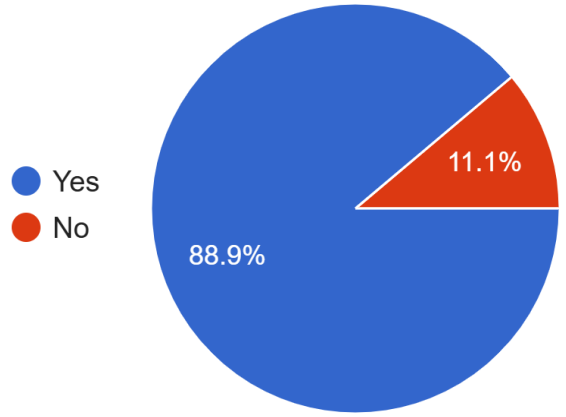


Topic of your big geospatial data project?

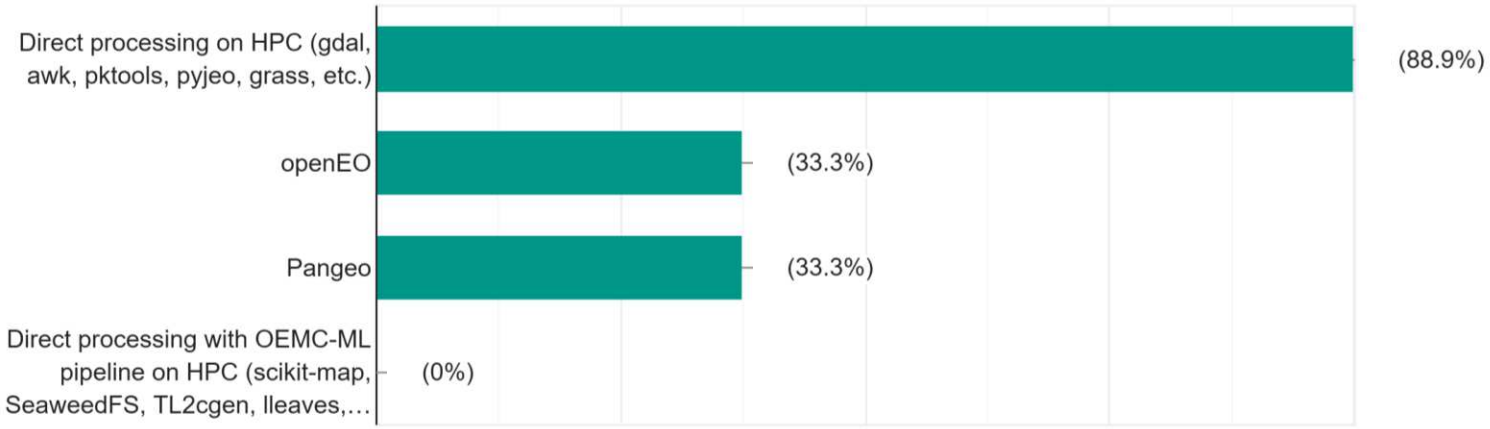


Open-source HPC Approaches

Will you be using open-source HPC approaches in your current and future projects?



Please select HPC approaches that you will use in the future?



Direct Processing on HPC

Understanding the architecture

Complexity of pktools and pyjeo

Complexity of apptainer

Software versioning & configuration

Terminal interface

Limited debugging information

openEO Platform

Complicated registration process

Performance issues

Backend navigation difficulties

Several reported no challenge

PANGEO

Complicated registration process

API-related issues

Complicated training assignment

OPEN EARTH MONITOR

Understanding the architecture

Complexity of SeaweedFS

Complexity of TL2cgen

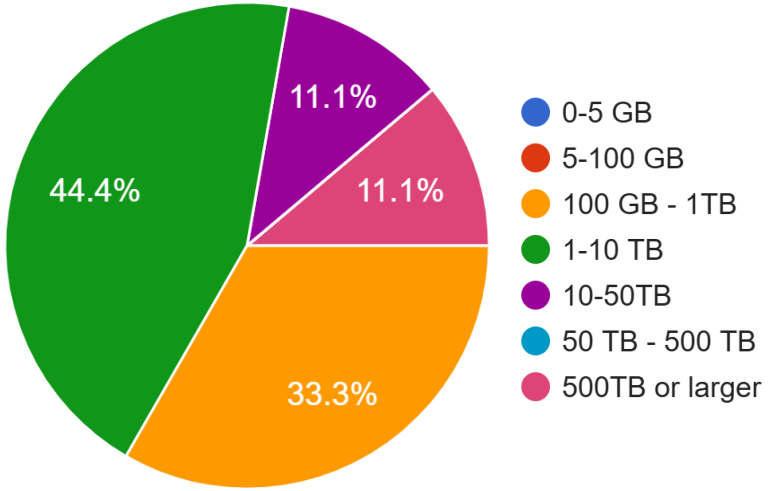
"hardcore"

Stability and versioning issues

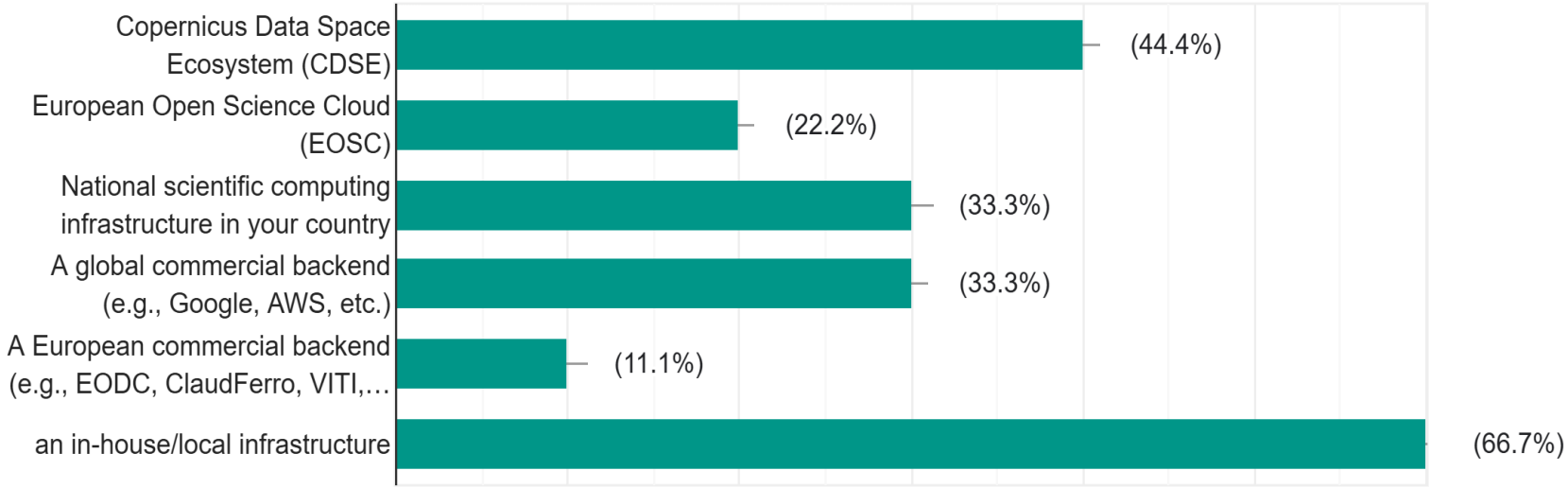
"I haven't covered it at all"

Data Volume & Backends

What data volume will you process in your next big-data geospatial project?

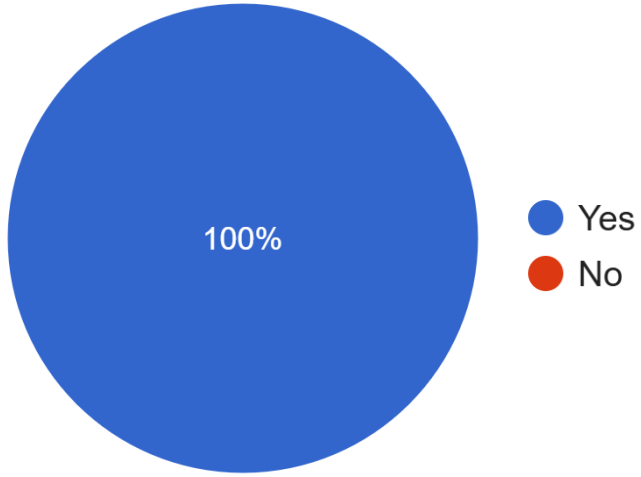


What backends will you use for your next big-data geospatial projects?

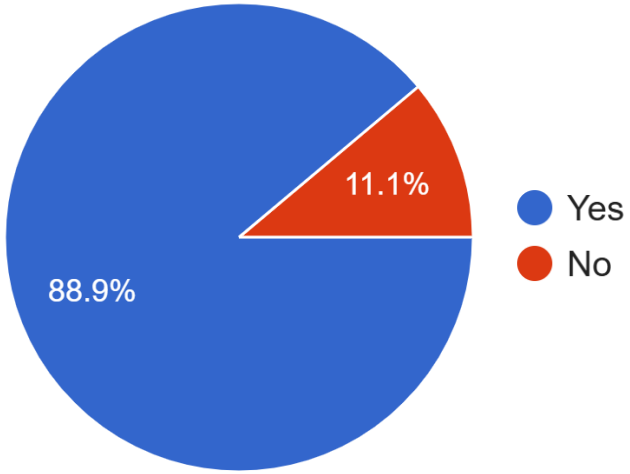


Open-source & HPC

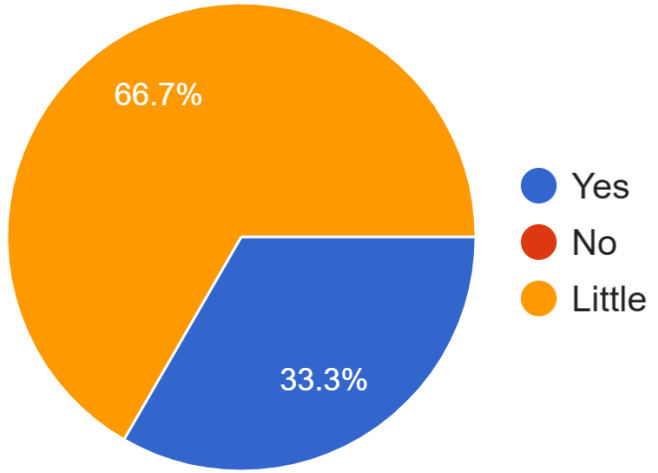
Are you interested in using open-source tools in the future?



Will you be using open-source HPC approaches in your current and future projects?

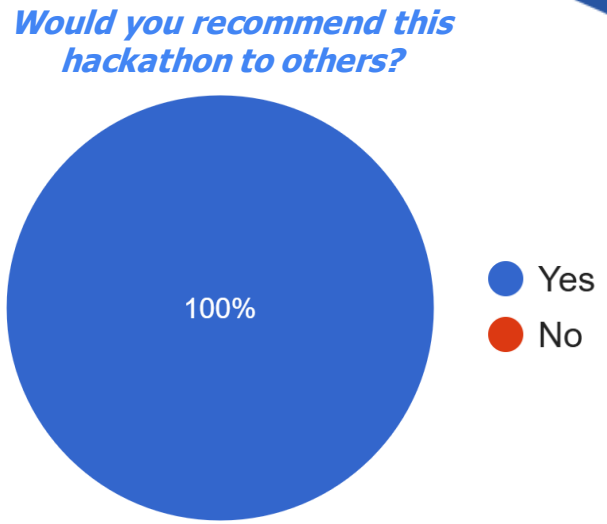


Do you feel confident in selecting the most suitable HPC approach and backend for your next big data geospatial project?



GEO-OPEN-HACK-2024

- Less than 3% participants from the industry!
- Typical big data projects:
  - most users still work with 10-100 GB of data that are locally stored
  - > 70% of users will have projects between 100GB and 10TB
  - most of users prefer to use local infrastructure or CDSE
  - users prefer direct processing on HPC the most
- Hackathon had a one-time budget



International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at

OPEN EARTH MONITOR

https://spatial-ecology.net/
SPATIAL ECOLOGY

SURF

openEO Platform

PANGEO
European Pangeo Community

OpenGeoHUB

Thank you!

milenkovic@iiasa.ac.at

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101059548.