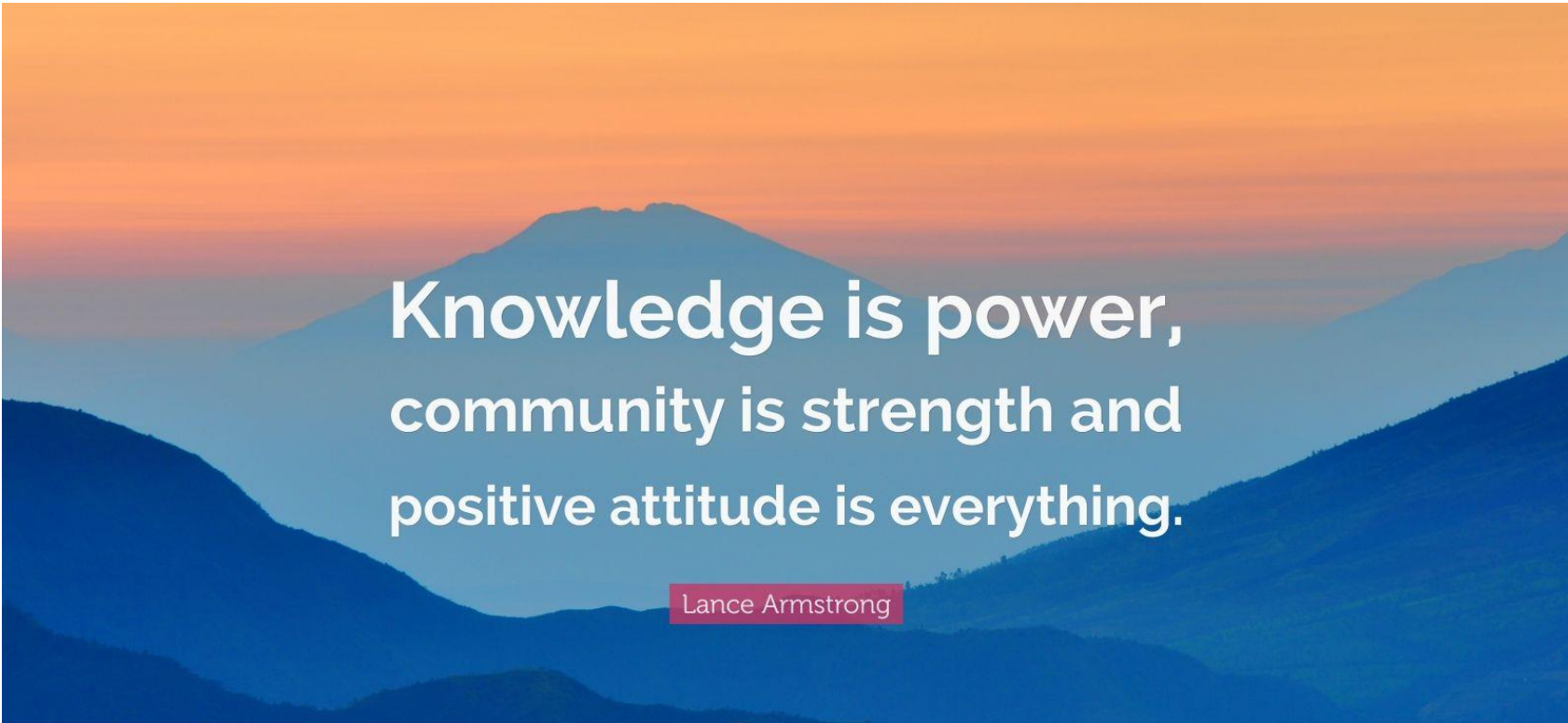




OPEN DATA & OPEN KNOWLEDGE Workshop

GEO Infrastructure Way Forward

Paola De Salvo
GEOSS Coordinator
GEO Secretariat



Knowledge is power,
community is strength and
positive attitude is everything.

Lance Armstrong

History: GEOSS Infrastructure Journey since GEO's Birth

....Together, the GEO community is creating a Global Earth Observation System of Systems (GEOSS) to better integrate observing systems and share data by connecting existing infrastructures using common standards. There are more than 400 million open data resources in GEOSS from more than 190 national and regional providers such as NASA and ESA; international organizations such as WMO and the commercial sector such as Digital Globe. ([About us \(earthobservations.org\)](http://earthobservations.org))



IIB

GEOSS Evolve

GD-07



5. GEOSS Knowledge Base Development

1 OVERVIEW



2nd GEO Data Providers workshop (20th-21st April 2017, Florence, Italy)



2

Days

38

Hours of
content

9

Sessions

8

Practical sessions

176

Presentations

95

Speakers

77

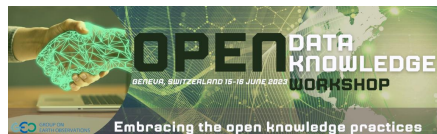
Organizations

193

In person attendees

~100

Online attendees



GEO Infrastructure Way Forward

Satellite

a) Establishing interoperability with cloud providers and their respective data catalogs;

b) Establishing interoperability with Data Cubes through the Open Earth Alliance (OEA) for increased uptake of the services and sharing of knowledge;

c) Strengthening interoperability with CEOS and Regional and National Space Agencies to allow users to access Analysis Ready Data (ARD);

d) Strengthening and establishing interoperability with regional/national services and data providers (in situ and satellite).



Data Cubes



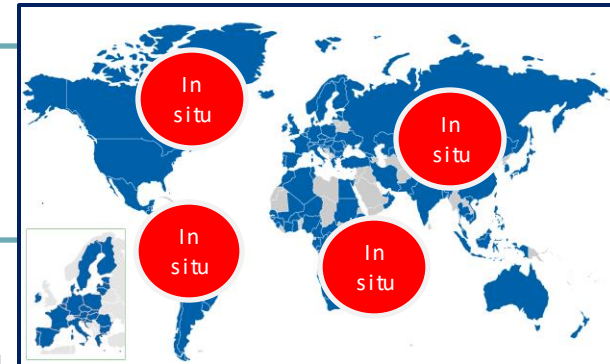
In situ

The uptake of in situ data, has been a key aspiration of GEO since 2005.

Strengthen collaboration with regional and national in situ data providers using standards that are both machine-to machine and human-readable.

For GWP activities and the rest of the GEO community to have access to more in situ data, GEO needs to **increase cooperation and interoperability with these networks in the evolving infrastructure.**

The work of the Data Working Group's In-situ Subgroup and its strategy will be strengthened by this supporting architecture of **better connected in situ networks**, especially with the ones that are currently not brokered in the GEOSS Platform via the GEO DAB.



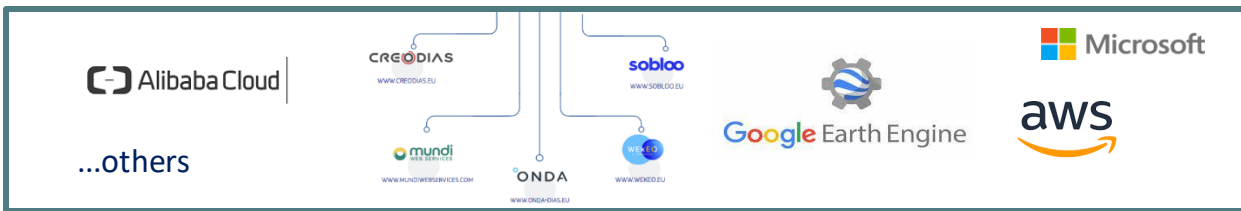
Technical analysis: Existing Components

Regarding existing components of the GEO Infrastructure:

- 1) Leverage the GEO DAB and its Application Programming Interfaces (APIs) to enable machine-to-machine use of data/knowledge;
- 2) Undertake the curation and quality control of current non space based data & services within the GEOSS Platform;
- 3) Integrate the current GEOSS Portal user interface with the GEO Knowledge Hub to provide a unique harmonized entry point for end users.



EAG Option 2



Data Cubes

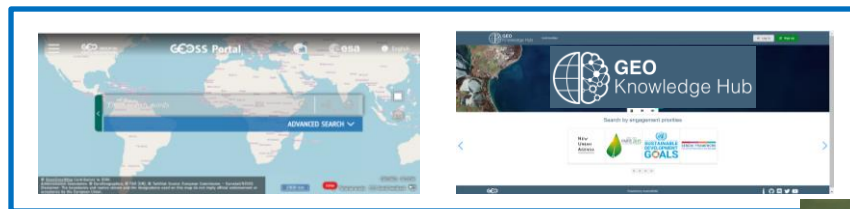


+ Commercial Companies (VHR)

EAG Option 3



Regional /National In Situ



Governance analysis:

Renewal of GIDTT membership is essential to make the evolution of the GEO Infrastructure more inclusive.

The renewed GIDTT will

- finalize any new architecture,
- identify resource needs,
- present proposal to ExCom in July 2024
- present implementation plan to GEO-20Plenary
- fall under proposed reformed Foundational Task “Data and Knowledge Management”.

GIDTT 2023-2025 will be inclusive and operational team

Thank You

Communicate and Collaborate with GEO





The project receives funding from the European Union's Horizon 2020 Research and innovation programme under grant agreement No 101039118

The GEOSS Platform Plus: the bridge from data to Knowledge

GEO Open Data & Knowledge Workshop, June 16, 2023

Gregory Giuliani (UNIGE), Mattia Santoro (CNR), Eliana Li Santi (SERC0 in support to ESA), Joost van Bemmelen (ESA), Paolo Mazzetti (CNR), Alessandro Scremin (Rhea in support to ESA), Daniele Giordani (RHEA), Marcin Trojan (EVERSIS), Małgorzata Herman (EVERSIS), Aleksandra Marczuk (EVERSIS)

GEO and GEOSS: the road from data to knowledge



Challenges



Solutions



Insufficient coordination among observation systems owners.



Insufficient data sharing (data often come at a price, especially until 2008).



Technological barriers due to data volume, heterogeneity, scattering, inadequate integration, accessibility and interoperability



GEO is established in 2005 an intergovernmental voluntary partnership (+100 Member states + EC, +100 organizations from different disciplines).



Fosters broad and open data



Builds the **Global Earth Observation System of Systems (GEOSS)** to improve availability, access, understanding and use of Earth Observations.

The GEOSS Platform: a bridge from data to Knowledge



Data, technologies, resources, services providers

International Data Providers*

Environment



Disasters



Biodiversity



Energy



Health



Urban



Water



Satellites



Food & Security



Africa



Brazil



Canada



Chile



China



France



Germany



India



Italy



Japan



Regional and National Providers*

Private Sector Providers



* a selection of more than 150 providers



Global Earth Observation System of Systems



FEDERAL GEOGRAPHIC DATA COMMITTEE



Status Checker



Yellow Pages



End-users (e.g. Decision Makers)



Intermediate users (e.g. Developers scientists)

UPSTREAM

MIDSTREAM

DOWNSTREAM



With support from
EVERSIS



The GEOSS Platform: a bridge from data to Knowledge in GPP



Challenges



Only data discovery and access, no other resources such as services, code, documents, tools, information, knowledge



No relationships between resources that would enable to retrace the journey leading to results



Set up of Community Portals and Views strongly depends on GEOSS Platform team support: this limits the engagement of new communities.



Obsolete *discovery download process publish* paradigm (not suitable for big datasets)



Solutions



Discovery of services, information, knowledge, tools, models, algorithms, papers



Semantic links among resources that would enable *reproducibility, replicability, reusability and robustness analysis*



Developing tools that enable **self-creation of community portals and views** to foster a broader engagement of new communities



Leveraging Cloud technologies supporting **multi-Cloud approach** for knowledge generation

Technologies around us

- Store and process big satellite data remotely
- Cloud-based platforms were developed specifically for accessing, visualizing, and processing EO data (Google Earth Engine, Copernicus DIAS, etc.)

Cloud/Multi-cloud



- Pre-processing datasets during the ingestion phase, providing Analysis Ready Data (ARD)
- Making processing for knowledge generation more efficient for the envisioned analysis

Data Cubes



- Enabling the networking of sensors and actuators promises a new era of in-situ data acquisition
- Potentially a new data deluge with new challenges on storing, accessing, and processing these new datasets

Internet Of Things



- Move data (pre-)processing close to the sensors can help to address IoT challenges
- Envisioning a Cloud Continuum supporting data processing

Edge Computing



- Boosted by the advancement of data-driven approaches based on Machine Learning (ML) and Deep Learning (DL)
- Promise to deeply affect EO data processing as other fields

AI/ML

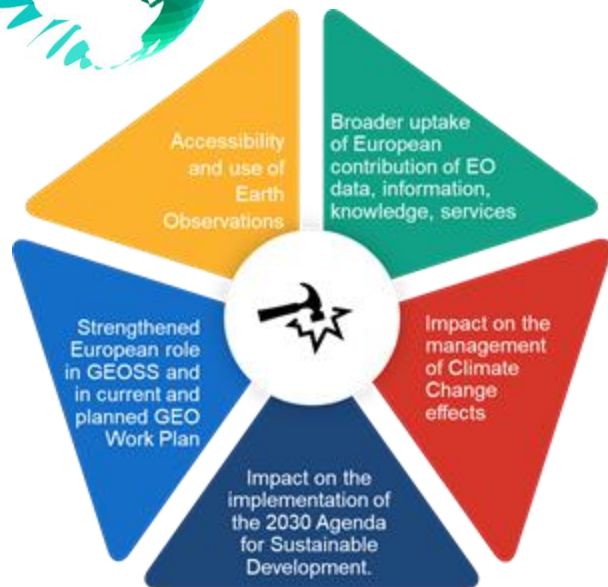


.....

*Each of these new technologies is a **potential enabler** for the desirable digital transformation of the process for **evidence-based environmental decision-making**.*

*However, it is also a **challenge for GEOSS** on how to exploit such a technology, and for **GEO** on understanding how the resulting innovation affects positively or negatively the transparency, reproducibility, replicability, and reusability of generated knowledge.*

Expected Impacts



GPP will further evolve the GEOSS infrastructure with users' required functionalities to ***access and generate tailor-made information & actionable knowledge***.

GPP will enable *services to non-specialists* in the domain of *adaptation to extreme climatic events* and to *changes in climatic conditions*.

It will implement different scenarios benefiting from GEOSS Platform developments. It will as well consider *linking with the GKH to set up the foundations of more interoperability among knowledge platforms*.

The GEOSS Platform: what exist now



A **data discovery and access system** bridging the gap between data providers and users.



One main HCI, the **GEOSS Portal**



Enabling the creation of **Community Portals** via customizable **GEOSS Mirrors**



Enabling the configuration of the domain of interest via customizable **GEOSS Views**



Accessible via open APIs, exposed by the **GEO Data and Access Broker** (middleware)



Replicability, Reproducibility, Reusability of experiments



In-situ repository



Data cubes



Knowledge sharing



SDG15.3.1 Dashboard

Reproducibility-replicability-reusability



SDG15.3.1

Community Portal Self creation tool

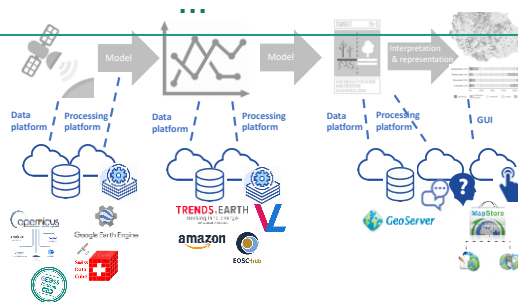
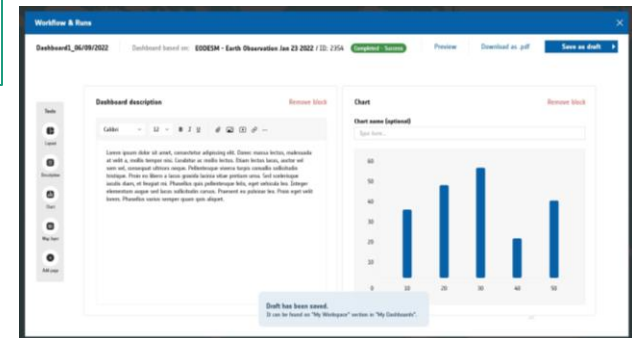
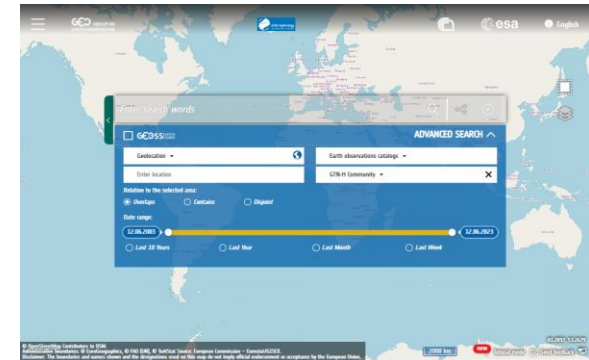


And enabling the configuration of the domain of interest via customizable **GEOSS Views**

Already selected:

1. Climate Change impact on Norovirus Pandemic Risk
2. SDG15.3.1 Land Degradation
3. SDG11.7 that relates to Climate Change, Urban Sustainability and Health, aimed to measure accessibility to urban green

Other scenarios to be integrated in 2023-2024, e.g. via Open ITT,



VLAB as enabler

Contribution and benefits

ESA ITT will be issued on July 2023

- Total budget available 500k for:
- Up to 250k for applications integration tasks
- Applications for deriving actionable information (climate change, GEOSS Platform enhancements)
- Not for development of new application but for the integration of Already existing applications

GEO Communities contribution:

- New Core Components developments: like moderators, harmonizers, Enablers (eg: GEO-DAB, VLAB)
- Application integrations and interoperability with the GEOSS Platform (eg: SDG15.3.1 and dashboard)
- Ad hoc Community portals features developments sharable with other communities (eg: GEO Community Portal Self Creation Tool, GEO Community portals enhancements)

Resource providers:

- New data repositories/catalogues registration (in-situ/model/other) : eg EEA in situ
- Cloud resources providers offerings
- Platform interfaces (eg: dedicated APIs / standards)
- Data cubes interfaces (Digital Earth Africa APIs)

The Benefits of contributing:

- Reproducibility, replicability, reuse
- Enhanced visibility and discoverability of own data, tools and resources in general;
- Single access point to heterogeneous (different source and nature) resources.

The proposed approach enhance:

1. *Reproducibility*: users can reproduce the experiment (same data/same analysis)
2. *Replicability*: users can replicate the experiment (different data /same analysis) >> use of national/local datasets instead of global ones.
3. *Reusability*: users can reuse/apply the approach in different contexts >> change the model and/or data sources.

GPP will propose possible evolutions of the GEOSS infrastructure with users' required functionalities to ***access tailor-made information & actionable knowledge***.

GPP will propose to enable *services to non-specialists* in the domain of *adaptation to extreme climatic events* and to *changes in climatic conditions* by allowing GEO Communities to co-design and interface their enablers with the platform, from which other communities can benefit.

It will demonstrate how to implement and interface different scenarios benefiting from GEOSS Platform developments. It will as well consider *linking with the GKH*.



*"In my lifetime, I've witnessed a terrible decline.
In yours, you could witness a wonderful recovery!"*

Sir David Attenborough, COP26 Summit, November 1, 2021