

Steps towards an In situ Data Coordination within the GEO Community

BACKGROUND

This concept note was written in collaboration between the GEO Secretariat and an ad-hoc team of people that had been active in the area of in situ coordination under the GEO umbrella in the previous years. The initialization of this work in early 2020 followed the Canberra 2019 Ministerial Declaration item 13¹ and responds, in a broader sense, to the need for action in this topic.

At the time this paper was first drafted, the terms of reference for the Data Working Group (DWG) had not been approved and no entity was officially in charge of tackling data sharing and management issues. ExCom-51 in March 2020 approved these terms of reference which describe duties to be carried out, some of them being particularly relevant for in situ data including the need to “Continue to promote the adoption and implementation of the GEOSS Data Sharing Principles and the GEOSS Data Management Principles globally” and “Advance discussions of critical data-related topics with stakeholder communities”².

The DWG Co-Chairs defined several main scopes of action including a dedicated subgroup for in situ. This paper serves as a basis for discussion and outlines possible future actions to be taken up by the GEO community.

¹ [Canberra Declaration](#), Canberra Ministerial Summit, November 2019

² Data Working Group: [Terms of Reference](#)

LIST OF CONTRIBUTORS

Ad-hoc team:

Carolina Adler -MRI, GEO GNOME

Henrik Steen Andersen -EEA

Stephan Dietrich - GTN-H

Yaniss Guigoz – UNEP/GRID Geneva, UNIGE

Wolfgang Grabs - GTN-H

Pierre Lacroix - UNEP/GRID Geneva, UNIGE

Michael Mirtl – ILTER, EAA

Hiroyuki Muraoka – ILTER, Gifu University

James Thronton – MRI, GEO GNOME

Data Working Group:

GEO Secretariat:

Douglas Cripe

Paola De Salvo

Florian Franziskakis

1 INTRODUCTION

Since its early days, the Group on Earth Observations (GEO) has recognized the importance of in situ measurements³ and the need to combine this type of data together with remote sensing imagery to generate greater impact. During the first decade of GEO, in situ coordination was the responsibility of the Work Plan Task IN-01 "Infrastructure".

However, in the face of a fragmented and rapidly evolving landscape, the GEO community has grappled with how to organize itself efficiently over this matter. Repeatedly, Declarations from GEO Ministerial Summits have called for strengthening this coordination (Cape Town 2007, Beijing 2010, Geneva 2014). The 2015 Mexico Ministerial Declaration outlined:

"(...) Earth observations take many forms including in situ, remotely sensed, and space based. Welcome GEO's inclusion of all of these types in its efforts."⁴

Launched in 2016, the GEOSS In Situ Observation Resources Task Team followed the work by the previous Task IN-01 and delivered a report in 2018 highlighting the need for coordination of in situ data and the potential benefits that could result from it.

Several recommendations stemming from their analysis called for: (i) development of sets of domain-specific Essential Variables (EVs) and identification of those EVs for which in situ data is required; (ii) regional-to-global coordination, where needed, of in situ data regarding standards and compliance with Global Earth Observation System of Systems (GEOSS) Data Sharing and Management Principles, aligned with the needs of the GEO Work Programme (GWP); and (iii) a gap analysis to understand where in situ data is still needed in the scope of the GWP as well as a guidance document for the maintenance of existing in situ networks.

Within the scope of this Task Team, an analysis revealed that, at the time, 63 networks of in situ measurements were deployed, 6 of them being dedicated to the atmosphere, 28 for the oceans, 25 on land and 4 were integrated ones.

Again in 2019, the Canberra Ministerial Declaration clearly recognized:

"(...) the critical role that data collected from the atmosphere, land and water (in situ data) plays in achieving GEO's mission, but note that sharing of such data is limited and there remain significant gaps in the global observing system. Call upon the GEO community to develop a strategy to address the challenges in this area and to demonstrate progress in implementation."¹

As the GEO community is moving towards a results-oriented strategy for GEOSS, taking steps to go from open data to open science while maintaining efforts to promote the use of Earth observations in decision making, it is critical to recognize that, in many areas, an efficient, organized, sustained coordination of in situ data is yet to be established.

³ Here, in situ refers to all land, water, and air-based observations, independent of the observing technology and methodology, excluding only space-based observations"

⁴ [Mexico City Declaration](#), Mexico City Ministerial Summit, November 2015

2 CHALLENGES

Several obstacles stand in the way of implementing efficient and sustainable coordination of in situ data. Some are technical but most of them relate to policy issues and are exacerbated by the complex landscape of GEO.

Technical challenges mainly concern the need for consistency among the many different observational networks collecting the data, as well as having sufficient/continuous spatial and temporal coverage in measurements.

The major challenge however pertains to following existing policies. In order to achieve results, the community must be able to use open and free data coming from authoritative sources, with complete metadata, clear guidelines (in particular in terms of formats), methodology and sufficient funding to ensure a long-term legacy.

Currently, researchers are hesitant to provide access to their in situ datasets because a) they could lose control over their research and may enter into a competition for results, and b) they may not receive adequate recognition for their work and contribution (data attribution).

3 GEO KEY PLAYERS

The very nature of GEO, being a best-efforts partnership between countries, organizations, and associates from the private sector, is adding a layer of complexity in the attempt to establish efficient coordination of an already fragmented landscape.

This section examines the potential role of GEO's key players namely (i) GWP Activities, (ii) Member States, (iii) Participating Organizations, (iv) Regional GEOs and, (v) Commercial Sector.

3.1 GWP ACTIVITIES

Activities of the GEO Work Programme are the primary instrument to facilitate collaboration and achieve GEO's Mission and Vision through coordinated efforts from all entities involved in GEO.

As such, these activities produce results based on Earth observations, whether coming from remote sensing imagery, in situ observations, or a combination of both. For the latter, in situ data is used for calibration and as an input for training and validating classification algorithms.

Frequently, however, activities of the GWP encounter challenges mainly related to issues such as accessing data and finding consistency when working with disaggregated sources. A major issue also relates to finding data with sufficient spatial and temporal coverage to create products of sufficient value.

3.2 MEMBER STATES

The condition for Member States to engage in GEO is that they endorse the GEO Strategic Plan “2016-2025: Implementing GEOSS”⁵, which includes the adoption and implementation of GEOSS Data Sharing and Management Principles.

Although progress on this implementation has been made over the last years and continues, broad, open data policies have yet to be implemented on a systematic basis, especially regarding sharing of in situ data.

As entities giving mandates and providing funds to collect data, Member States should be aware of the value of open data sharing and the opportunities it generates (both socially and economically). Hence, it is critical that over the next years, efficient policies addressing open data sharing be put in place in countries where this is currently not the case.

Countries should see the implementation of such policies as an investment, not as a potential threat. Having data locked in a drawer brings very little value to it, whereas making it available for anyone to use only enhances its potential use by many actors including other national public services, NGOs, academic researchers, international organizations as well as the commercial sector.

3.3 PARTICIPATING ORGANIZATIONS (POs)

As the term "Participating Organizations" encompasses a variety of agencies, research institutions, networks and so on, it would be difficult to draw a complete picture. Yet some common action items could be envisaged from POs.

As data providers, brokers of relevant data in their given domains, and custodian agencies, POs in GEO could play a crucial role in advocating open data sharing at their respective scales (local to national or international).

POs could also continue the work on data interoperability and keep on putting efforts in harmonizing definitions of common parameters used in decision making and for the identification and the integration of data from different sources, such as EVs.

3.4 REGIONAL GEOs

As an intergovernmental partnership, GEO must deal with the different working arrangements of its Member States and thus faces the challenges of finding ways to collaborate solely based on voluntary contributions. It is often indeed difficult to establish harmonized practices at the continental scale. For this reason, Regional GEOS were established to conduct activities and coordinate efforts in a more proximity-based, inclusive, co-designed way as solutions to challenges that are common to the region are sought.

⁵ https://www.earthobservations.org/documents/GEO_Strategic_Plan_2016_2025_Implementing_GEOSS.pdf

Continental coordination dealing with in situ data collection, management, storage and sharing could be the most efficient way to address many current challenges, the main being adoption of open data policies and ensuring they are followed by data providers.

Collection, update and long-term preservation of in situ data organized on a continental scale could also be an efficient way to guarantee data quality and would allow an easy identification of gaps and further planning in order to obtain better coverage (spatially and temporally).

Regional GEOS have a crucial role to play as intermediaries between GEO and its Member States, especially in making sure GWP activities can benefit the latter. In this sense, a lot could be gained by stronger involvement of Regional GEOS in the in situ data related challenges.

3.5 COMMERCIAL SECTOR

Private companies, SMMEs and international groups active in Earth observations are legion, whether it comes to collecting data, creating algorithms to process it or to build integrated frameworks and services. Established in 2019, the "GEO Associate" category, now allows partners from the commercial sector to be formally involved in, and contribute to, GEO.

The role of the commercial sector regarding in situ data is quite complex to explore, especially through the optics of data sharing and management principles that clearly state all products should be made openly and fully available to anyone when working with GEO.

4 EXTERNAL PLAYERS

Despite the broad outreach of GEO, much of the existing in situ datasets is being collected within the frame of small to medium scale projects that have no direct link to GEO making it arduous for others to find and reuse it.

Everyone – not only GEO – would benefit from wide, organized and sustained sharing of data to report on common policy engagements. Engaging academia, research networks and other relevant actors to adopt an open data management strategy is a broad topic going way beyond the scope of this paper, however GEO can play a role in promoting an open science approach.

An example of such a strategy has been implemented with the various GEO cloud credits programmes, as all components of the projects (including in situ data) must be made freely available using a Findable, Accessible, Interoperable and Reusable (FAIR)⁶ compliant license.

⁶ Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>

5 NEXT STEPS

As described above, attempts to address in situ related issues in a coordinated fashion where not already happening must encompass all key players of GEO. It is crucial to note that these key players may in practice have more "fluid" roles than what is described above (e.g. a PO using a GWP's products, GWP activities collaborating with the private sector in the framework of a Regional GEO's specific partnership etc....).

The 2020 - 2022 GEO Work Programme Foundational Task (FT) "GEOSS Data, Information and Knowledge Resources" contains a dedicated component for in situ observations aiming to:

"Improve access to in situ data and provide coordination and repository services where needed. Advocate new data collection systems and encourage integration of user requirements with respect to in situ data, which will be closely linked with the GEO Knowledge Hub."⁷

In the scope of this FT, several actions could be envisaged, including the following:

- **Focus** on a few GWP activities to identify their needs for in situ data and where possible, engage custodian agencies or any other provider to share them;
- **Develop** pilots to showcase the benefits of in situ data (to the research community in particular) to generate greater impact;
- **Ensure**, to the extent that in situ data forms part of a particular 'knowledge package' feature in the GEO Knowledge Hub, that the data referenced are properly identified (e.g. Digital Object Identifier), openly available and accessible;
- **Promote** sustained and consistent new in situ data collection on land, air and water, based on requirements from satellites (especially future projects through which in situ data collection should be defined and funded);
- **Discuss** the importance of trusted repositories for in situ data at regional to national scale, as well as the need for data life cycle management;
- **Assess** the legacy of in situ data collected from research projects after their completion when linked to GEO.
- **Advocate** a larger role for Regional GEOs towards promoting open data sharing, based on GEOSS Data Sharing and Management Principles;
- **Investigate** the role of the GEOSS Infrastructure by inventorying the accessible existing in situ datasets and by evaluating if it can act as an hosting platform for in situ;
- **Explore** new technologies and solutions allowing use of in situ data by lowering the processing steps (in situ ARD) or when direct access remains restricted (e.g. machine-to-machine access for algorithm training purposes);

⁷ https://www.earthobservations.org/documents/gwp20_22/gwp2020_summary_document.pdf