Lessons Learned and Transition Plans for GEO Cloud Computing Programmes

This document is submitted by the GEO Secretariat to the Programme Board for discussion.

1 INTRODUCTION

GEO's mission is to enable the use of Earth observations (EO) to support sustainable development policies that are grounded on objective data and methods. In support of its mission, the GEO Secretariat is promoting the use of cloud services as the primary method for effectively mining information from satellite imagery big data. Cloud services allow users to focus on the production of information and analysis, rather than be distracted with the data management overhead, since cloud services minimise the efforts required to download, store, and manage large datasets.

In the spirit of exploring engagement with the commercial sector as directed by GEO Plenary (see Section 2 below), the GEO Secretariat negotiated an arrangement with Amazon Web Services (AWS) in late 2018 in which AWS cloud credits would be offered to successful applicants taking part in a competitive proposal submission. The GEO-AWS EO Cloud Credits Programme offered developing countries access to free cloud services to help with the hosting, processing, and analysis of big EO data for sustainable development. In particular, the credits being made available were intended to enable government agencies and research institutions to build EO applications that support environmental and development goals, including the Sendai Framework for Disaster Risk Reduction, the Paris Agreement and the United Nations Sustainable Development Goals (SDGs). The GEO Secretariat and members of the Programme Board served as evaluators of the proposals submitted to select the finalists, based on scientific excellence and a credible implementation plan, and the projects were launched in the summer of 2019.

Building on this initial arrangement with AWS, the GEO Secretariat issued a broader *Commercial Sector Engagement Opportunity* in September 2019 to invite proposals from companies that provide cloud-based geospatial processing platforms for development of applications using Earth observations. As a result, over the past year, the Secretariat has successfully launched two additional programmes to provide licenses, grants, technical and financial support to the GEO community, following a similar competitive call for proposals and combined Secretariat/Programme Board review process to select the winning applicants. As of this writing, GEO is actively supporting 57 projects (20 with Amazon Web Services, 32 with Google Earth Engine, 5 with Microsoft AI for Earth), in order to:

- Support projects in developing countries;

- Help GEO Members deliver on global policy mandates; and
- Tackle planetary issues with open EO data.

2 WHAT WAS THE MOTIVATION FOR THE GEO SECRETARIAT TO ENGAGE WITH COMMERCIAL PROVIDERS[1] OF CLOUD SERVICES AND ANALYTIC TOOLS?

The *GEO Strategic Plan 2016-2025: Implementing GEOSS,* endorsed by GEO at the GEO-XII Plenary held in Mexico City (2015), notes that the strengths of GEO lie in its distinctive characteristics, including being "a flexible and agile forum for governments, public sector agencies, UN bodies, specialized organizations, universities **and the private sector** to work together on improving the quality, timeliness, range and availability of Earth observations, data, information and knowledge about the Earth system."

The *Strategic Plan* went on to devote a section to outlining potentially mutual benefits resulting from engagement with the private sector:

> *"GEO offers unique information and engagement opportunities to the **private sector** and international development banks, to serve their needs in areas such as agriculture, transportation, resource extraction and insurance sectors. In turn, GEO benefits from the participation of the **private sector** through access to new types of data, diverse capabilities and new technologies, and broader community networks. […] Given these mutual benefits, the **private sector** and development banks have the capacity to be key contributors to making substantive progress towards achieving GEO's Strategic Objectives."*

The *Strategic Plan* also suggested basic criteria and milestones that should characterize GEO's interaction with the private sector:

> *"A sound, sustainable financial footing is required for GEO to be successful through 2025. Any mechanism to encourage contributions will respect the voluntary nature of GEO, be simple, adaptable, and applicable to GEO's context. The mechanism should also foster greater engagement with the **private sector** and philanthropic foundations while maintaining coherence and cohesiveness across GEO."*
>
> *"As GEO enters the next decade it will focus on addressing the societal challenges facing humankind by being an advocate for the value of Earth observations, engaging with key stakeholders including the **private sector** and development banks, and delivering data, information and knowledge critical to informing decision making."*

Finally, the Mexico City Declaration called for GEO to: "develop new approaches to effectively engage with: United Nations institutions, Multilateral Environmental Agreements, Multilateral Development Banks, additional Participating Organizations **and the private sector.**"

[1] In the discussions that follow, notwithstanding the finer differences in nuance that may be marked by different communities, "commercial" and "private" are used interchangeably to signify services provided by for-profit businesses and entities.

Based on the Strategic Plan and the Mexico City Declaration, with guidance from the Executive Committee, the Secretariat drafted an Engagement Strategy that was approved a year later by the GEO-XIII Plenary in Saint Petersburg (2016). The Engagement Strategy included Section 4.5 *Private sector (commercial entities and private users)* which recalled that:

> *"At GEO-IX in 2012, the GEO community recognised the value of engaging with the* ***commercial sector****, while still protecting the organisation's values and purposes" and notes "the* ***commercial sector*** *can contribute to jointly identifying user needs while benefiting from increased data sharing. Benefits can also be drawn from the exploitation of integrated Earth observations to develop commercial applications, tools and other products that address specific user needs and enhance decision-making."*

In addition, the GEO-XIII Plenary agenda featured a "Dialogue with the Commercial Sector" session which explored challenges and opportunities related to Commercial Sector engagement in GEO, from contributing or benefiting from existing GEO efforts to exploring opportunities for new collaborations. It was also noted that the GEO Discovery and Access Broker (GEO DAB) brokering framework, part of GEOSS implementation, is making use of several facilities available through Amazon Web Services, such as cloud Logging as a Service (LaaS) and Platform as a Service (PaaS) as well as load balancing, DNS routing, auto-scaling, monitoring, elastic map reduce, storage, and computing. This example provided a potential model for further GEO engagement with the private sector.

Along with GEO-XIII Plenary's approval of the *Engagement Strategy*, the *GEO Rules of Procedure* were modified with the addition of an initial version of *Annex C* ("Engaging the Private Sector in the Implementation of the Global Earth Observation System of Systems GEOSS"). *Annex C* represented a first attempt to lay out in detail the principles and modes of engagement that should govern engagement with the Private Sector. The next update of the GEO Rules of Procedure, approved by GEO-XIV Plenary, included a greatly revised and expanded *Annex C* (now entitled: "Rules of Engagement with the Commercial Sector"). Subsequent Plenary approvals of updates to the GEO Rules of Procedure included additional minor refinements to *Annex C,* placing an increased emphasis on transparency and openness of any engagement processes with the commercial sector.

Accordingly, the *Commercial Sector Engagement Opportunity* announcement stipulated that: "offers from commercial providers must abide by Rules of Engagement with the Commercial Sector contained in Annex C of the GEO Rules of Procedure, in particular articles 2.6 and 2.7". Further, the announcement noted that this was an ongoing effort by the Secretariat in which:

> *"Any commercial sector cloud services provider may respond to this engagement opportunity. There will be no competitive selection of commercial providers. In principle, all offers are welcome and will be considered on their individual merits. Commercial sector cloud services providers are welcome to make proposals to the GEO Secretariat with respect to grants and/or credits for use of their services at any time."*

3 WHAT ARE THE HIGHLIGHTS OF THE GRANTS AND LICENSE PROGRAMMES?

- The GEO AWS programme provides USD$ 1.5 million in credits for use of AWS over a 3-year period. An additional € 300K credits from Sinergise to use Sentinel Hub were made available to successful proposals. 20 projects were selected for credit awards (from a total of 30 submissions).

The GEO-GEE programme provides commercial-grade license for unlimited use of GEE over a 2-year period, valued at over USD$ 3 million. This arrangement is of benefit even to research institutes as there are no data quotas and GEE can be used on an operational basis. 32 projects selected for license awards (from a total of 52 submissions).

- The GEO-Microsoft programme include $500K in Azure cloud credits and $500K in funding over 1-yr period for GEO BON projects ($100K max per project). 5 projects were selected for funding and cloud credits (from a total of 60 submissions).
- Results will be featured at major GEO events, such as side-events during GEO Week, as well as in blog and news articles.
- There has been a broad geographical distribution regarding institutes submitting proposals to these programmes, across both developed and developing countries.
- Participation in the programmes have saved countries' time and energy in regard to dealing with administrative issues.
- In some cases, the programmes provided access to grants where none previously existed (e.g. Iran, Iraq and the Ukraine).
- These grants were negotiated in the spirit of GEO, i.e. according to the GEO Rules of Procedure.

4 LESSONS LEARNED

Since the GEO-GEE license and GEO BON-Microsoft programmes are only a few months old, these remarks pertain to the GEO-AWS EO cloud credits programme in general.

4.1 The Good

- The review process, comprising members from the GEO Secretariat and the Programme Board worked smoothly and efficiently, allowing an impartial selection of the best proposals from all those submitted.
- A number of the projects are moving along rapidly with impressive results after only one year, demonstrating that it is feasible for developing countries to successfully leverage big EO data and cloud computing platforms for application development. These projects are incredibly appreciative of the opportunity afforded by the GEO-AWS programme. For example:

"At the current research phase of the project, access to AWS cloud infrastructure and services is critical because our institutional capabilities do not allow us to host and administer the necessary environment on our own." Oleg Seliverstov, Illegal amber mining project, Ukraine

- Having the Sentinel Hub credits from Sinergise allowed the Iran/Iraq proposal to move ahead quickly.

4.2 The Bad

- It became evident early on that targeted training materials would need to be provided by AWS to the groups.
    - An obvious consequence of aiming for this particular target group (developing countries) was that many simply do not have the capability to start up on the cloud. This has led to underutilization of credits as some groups who were not fully prepared to work on AWS. The implication is the need for a better technical support framework from the beginning.
    - Currently, there is a lack of open-source tools to help countries explore big EO data.
    - AWS had assumed there would be an opportunity to meet with groups when already in a central location (GEO event, industry meeting, etc) to run training sessions, but this did not appear feasible for these groups given the expense of international travel (COVID-19 notwithstanding).

  Remedy: A discussion forum has now been set up where capacity development related to technical and other issues can be addressed more quickly and systematically.

- Despite initial efforts, the groups are all starting work at various times. In the (hopefully) most extreme example, one group is starting one year into the programme. The reasons for this ranged from changes in personnel, being out on fieldwork, and general communication difficulties in some developing countries.
    - As a consequence, it has taken longer than expected for AWS to connect the grantees to the full range of account resources. Because all groups are at different stages it has been a challenge to complete this administrative phase uniformly.

  Remedy: This imbalance should even out as year two progresses.

4.3 The Ugly

- One of the largest awardees (INEGI, Mexico) revealed after the fact that, for legal reasons, they are currently unable to directly accept or utilize the AWS credits. It should have been made more explicit at the outset that awardees agree to Terms and Conditions ***prior*** to applying for the programme.

  Remedy: INEGI and AWS have been in contact to try to find a solution. In the meantime, a work-around is being implemented through one of the consortium partners located in Colombia and the project is moving forward.

5 HOW HAS PARTNERING WITH THE PRIVATE SECTOR BROUGHT VALUE TO THE GLOBAL GEO COMMUNITY? (MICROSOFT AI FOR EARTH, GOOGLE EARTH ENGINE AND AMAZON WEB SERVICES)

The "zero download model" (i.e. no EO data or analytical software downloaded locally) has been successfully demonstrated, lowering barriers to accessing and analysing EO data:

- Since data are shared in a public cloud computing environment, researchers no longer have to worry about downloading or copying data before getting to work; and

- Data users can analyse massive amounts of data in minutes, regardless of where they are in the world or how much local storage space or computing capacity they can access.

Multiple benefits of the open science approach have also been demonstrated:

- Having an open, common set of data to work from brings consistency to results;
- Risk of erroneous or "stale" data reduced due to increased scrutiny;
- Cross-collaboration is encouraged and serves as a reality check on many levels; and
- Open access to data and computing environments enables broader uptake (especially for developing countries).

Additionally, EO applications are made scalable, reliable, and robust by means of stable, global cloud platforms at minimal cost and infrastructure investment for developing countries. Also, opportunities for engagement created through GEO Week Industry Track provides the chance to learn from global industry players and their millions of customers feeding back into the product.

6 HOW WILL THESE INITIATIVES LINK TO THE GEO WORK PROGRAMME?

Some of the projects stem directly from Flagships or Initiatives of the GEO Work Programme. Examples include:

- **GEO BON** has been developing the concept of Essential Biodiversity Variables (EBVs), which are a minimum set of biological state variables that are needed to detect changes in biodiversity. However, it has proven difficult to produce EBVs on an operational scale, due to challenges of integrating in situ and remotely sensed Earth observation data within a single geo-computation environment, as well as a host of other issues related to limitations in accessible computing resources, such as workflow execution, and product storage and provision. The project under GEO-GEE will scale-up a suite of complementary EBV data products to application -ready time-series products and deliver these as open access products. The wider GEO BON network will be engaged to ensure that the developed EBV data products can feed into indicators for ongoing and upcoming policy processes.
- **AquaWatch** is: a) working to obtain scientific community consensus for an operational system design for satellite remote sensing of surface water quality addressing both atmospheric correction and water quality algorithms; b) generating chlorophyll-a, total suspended solids and coloured dissolved organic matter (CDOM) products using Google Earth Engine; and c) ultimately constructing a near real time (within constraints of data streams), interactive portal of surface water quality with global coverage.

In other cases, projects are just getting underway and the Secretariat will be looking for ways to connect them to appropriate GWP activities, especially those dealing with some aspect of helping GEO Member countries report on their commitments under global policy agreements, such as the SDGs, Paris Agreement and Sendai Framework. As an example, the Brazilian AWS project is already producing rapid response modules that can be integrated into Global Forest Observation Initiative (GFOI) Flagship and Digital Earth Africa algorithms for forest monitoring.

7 HOW HAVE ISSUES AROUND INTELLECTUAL PROPERTY, DATA OWNERSHIP AND OPEN DATA BEEN APPROACHED?

Article 3.4 on Ethical Standards and Intellectual Property from the GEO Rules of Procedure states:

> *"Data, software, products, and other goods developed solely or jointly by GEO Members, Participating Organizations and Associates during and within the scope of a collaboration within a GEO Work Programme Activity should not be made subject to copyright or intellectual property right restrictions, but should be made available to the public under an open license, at no cost and with no restrictions on copying, publishing, distributing, citing, or adapting."*

As a consequence:

- In all cloud credits programmes launched by GEO in partnership with commercial companies, it is stipulated that the credits are provided to the institutions subject to the conditions that all data and software developed in the course of the project will be made available globally as open source.
- All data used in the project are to be fully documented (showing provenance and access conditions, if any), and made available in a free and open manner, according to the GEOSS Data Sharing Principles and the FAIR[2] guidelines.
- Data from third party space-borne platforms used in the project (either in its original or processed formats) should be made available using the same license and giving the same rights of use as the original data.
- Data from in-situ measurements provided by third parties used in the project should be made available using the same license as the original data.
- Data from in-situ measurements produced in the project should be made available using a FAIR-compliant data license, such as Creative Commons or the Open Database License.
- All software used to produce the results, including the base packages and user-developed algorithms, should be made available using one of the licenses approved by the Open Source Initiative such as the GNU General Public License.
- The GEO-AWS cloud credits programme has recently completed year 1 (out of 3). As most of the projects are still in their initial phases, the number of software products currently available is limited. Nevertheless, all awardees have confirmed their intention to proceed with their plans and comply with their obligations regarding open software availability.

8 WHAT HAPPENS AT THE END OF THE PROJECT PERIOD?

No organization is under any requirement to continue using any of the platforms or the cloud services that they have been working on. However, it would be reasonable to expect that, if a given project has demonstrated beneficial results, that the proponents enter into commercial discussions for long-term support.

[2] Findable, Accessible, Interoperable, Reproducible.

There may be options to extend the terms of the grants or even possibilities to work with some of the major funding bodies who can support these activities in light of the important work in the areas of biodiversity, climate change, disaster risk reduction, and so forth. In the context of the programme duration, the institutions involved can assess the value of using commercial cloud providers for the period before making decisions on whether to buy services for them, since:

- The value of the grants is significant and allows institutions to use cloud services for realistic projects.
- The institutions involved in the programmes are developing open source, good quality software which will ultimately be shared with the GEO community.
- Collaborative open source development can provide substantial cost savings when using cloud computing (i.e. benefits outweigh costs).

In some cases, the cloud credits programme is being leveraged to develop a stand-alone tool or data analysis platform that can then be downloaded locally and shared with other national institutions. This is the case with the Brazilian data cube project under AWS, in which the credits are being used by the chief institution (INPE) to process and build a data cube for analysis of Brazilian EO. At the end of the project, the data cube will be available both on AWS and as a standalone product that can be made available to other Brazilian public institutions, providing an environment to explore EO data without using AWS.

Beyond this, the GEO Secretariat is exploring alternatives with respect to sustainability. As an example, for applications that involve forest monitoring, the Secretariat is discussing with the GFOI Flagship the use of Food and Agriculture Organization (FAO)'s System for Earth Observation Data Access, Processing and Analysis for Land Monitoring (SEPAL) cloud-based computational platform, that allows GFOI-linked institutions to run software in both Google Earth Engine and Amazon Web Services, at no cost to the institution. SEPAL is funded by the Government of Norway. The Secretariat considers that the example of SEPAL is a noteworthy one, where financial resources from donor countries can be concentrated for more effective use. Another option would be to embed successful applications and products into larger initiatives that have already assembled resources for development. One example would be the Digital Earth Africa project; in cases such as these, the success of the project could be expedited by leveraging algorithms from the cloud credits programmes that have already been developed and proven to be effective at processing big EO data in the data cube format.

9 REPORTS FROM AWS PROJECTS.

First-year reports from several projects under the GEO-AWS EO cloud credits programme may be viewed here:

https://drive.google.com/drive/folders/1jqqa6SjaSuAz3Qtq76uJnIwJSCrtlmJP?usp=sharing

These preliminary reports will be supplemented by a more detailed summary by the end of 2020.