

1. INTRODUCTION

The Group on Earth Observations (GEO) Knowledge Hub (GKH) (<https://geo-knowledge-hub.org/>) is a digital repository providing access to the knowledge required to build and reproduce applications of Earth observations (EO). The purpose of the GKH is to reveal all components of a given application produced in an open science context for replication and up-scaling, including: (a) research papers and reports describing methods and results; (b) software algorithms and cloud computing resources used for processing; (c) in situ and satellite imagery data used; and (d) results for verification.

2. WHY IS THE GEO KNOWLEDGE HUB NEEDED?

GEO has the mission of improving the capacity of all its Member States to use Earth observation data for decision-making. The GEO community produces many useful and relevant results, which GEO aims to share globally. Achieving this goal requires broadening global access to the knowledge produced within GEO, as well as combining best practices from the GEO community with long-term capacity development. The GKH will lower the barriers for developing countries in particular to use the petabytes of big Earth observation data openly available.

3. WHO IS IT FOR?

The GKH will be useful to a wide range of stakeholders, from national experts needing to report on policy commitments, to individual end users and Small- and Medium-sized Enterprises (SMEs) seeking practical solutions to local environmental challenges. Activities of the GEO Work Programme (GWP) will be fundamental in both providing methodologies for solving problems and identifying potential end users. Applications derived from the GWP such as the Arctic GEOSS Community Activity can be made available for scaling up from local and national to regional and global levels by means of the GKH. Technical experts from research institutions may serve as intermediaries in assisting local end users to benefit from the resources of the GKH. GEO intends to leverage the capacity development networks of its partners in a “training the trainers” approach.

4. WHAT RESOURCES WILL IT CONTAIN?

The contents of the GKH are linked documents that contain relevant information for Earth observation applications that promote reproducibility, scalability, and co-design/co-production. Examples of documents include an HTML file, a PDF file (report of paper), a Jupyter Notebook, an R or python markdown file, a GitHub page, a repository entry linking to a dataset store with an assigned Digital Object Identifier (DOI), a suitable cloud computing environment, and links to other resources such as instructional videos.

With the GKH, users will have a single-entry point to discover and access resources that have been developed by domain experts. Such resources will be of different types to reflect the entire information flow of the research and knowledge pertaining to a domain. Table 1 below provides a summary of these resources.

Table 1 – Types of knowledge resources organized in the GEO Knowledge Hub

Knowledge resource types	Sources
--------------------------	---------

<i>Publications</i>	Scientific Journals and Reports provided by GWP activities, with associated unique digital object identifier (DOI). Post-print copies will be stored in the GKH.
<i>Software code, models and tools</i>	Open-source software code provided by GWP activities, preferably available in Github, with associated DOI. Backup copies stored in the GKH.
<i>Remote sensing data</i>	Link to files in a cloud repository.
<i>In situ data</i>	Link to files in a recommended data repository, with associated DOI.
<i>Ancillary data</i>	Link to files in a recommended data repository, with associated DOI.
<i>Output data and products</i>	Link to files in a recommended data repository, with associated DOI.
<i>Videos</i>	Directly stored in the GKH (preferably).
<i>Other relevant documents (e.g., training material)</i>	Directly stored in the GKH (preferably).

5. HOW DOES IT WORK?

The GKH will typically be populated by knowledge resources developed through activities of the GWP, and involves joint work by the GEO Secretariat and the GKH contributors to ensure inclusion of trusted information. Since the GEO community needs best practices that are reproducible, information ingested into the GKH must be verified and organized, while disparate components of a document set must be linked. The GKH team at the GEO Secretariat will interact with the contributors so that the methods, data and software are consistent and follow the GEO Data Sharing principles, the GEO Data Management Principles and other applicable Open Science principles.

EO application ingestion into the GKH includes the following actions:

- Identify EO applications, developed following an open science approach, suitable for ingestion. Potential contributors will be contacted by the GEO Secretariat to request shareable copies of knowledge resources, including methods, software, and data.
- Contributors will be asked to develop best practices for promotion of sharing, such as depositing in situ data on trusted repositories, making software available in GitHub (or similar platforms) and assigning DOIs to software and to documents where needed.
- Promote the use of cloud computing solutions (encourage contributors to provide versions of their results that work in cloud computing environments).
- Enter information about the links that connect the publication, software and data into the indexing system of the GKH.
- All issues of data quality and software stability/reproducibility are the responsibility of the contributors.

The ingestion process of the GKH will be as automated as possible, but human intervention may be required for final verification.