

Linking Science and Technology to GEOSS: Contributions of the ST-09-02/ID-03 GEO Work Plan Tasks and the EGIDA Project

Hans-Peter Plag¹⁾ and Ian McCallum²⁾

¹⁾IEEE and University of Nevada, Reno, Nevada Bureau of Mines and Geology, Reno, Nevada, USA, hpplag@unr.edu; ²⁾International Institute for Applied Systems Analysis, Laxenburg, Austria, mccallum@iiasa.ac.at



GEOSS AND SCIENCE AND TECHNOLOGY

The Global Earth Observation System of Systems (GEOSS) developed by the Group on Earth Observations (GEO) aims to provide comprehensible Earth observations (EOs) in support of decision making in a wide range of societal benefit areas. The nine interdependent Societal Benefit Areas (SBAs) addressed by GEO require an interdisciplinary scientific approach, and scientific interpretation of the EOs provided by GEOSS is necessary in order to derive actionable information. A strong engagement of science and technology (S&T) communities in both the development and use of GEOSS is necessary to address the complex issues of the global integrated Earth system; improve interoperability between global observing, modeling, and information systems; facilitate data sharing, archiving, dissemination, and reanalysis; optimize the recording of observations, assimilation of data into models, and generation of data products; enhance the value of observations from individual observing systems through their integration in the SBAs; and harmonize well-calibrated, highly accurate, stable, sustained in-situ and satellite observations of the same variable recorded by different sensors and different agencies.

OUTREACH TO S&T COMMUNITIES

The S&T Road Map of GEO defines a number of outreach activities in order to ensure a broad S&T support for GEOSS and a wide use of GEOSS products by S&T communities. The 2009-2011 Work Plan Task ST-09-02 (see <http://www.geo-tasks.org/st0902>) has implemented these activities, and the new 2012-2015 Work Plan Task ID-03 will continue to develop them. Since 2010, ST-09-02 has been supported by the EC-funded EGIDA Project (<http://www.egida-project.eu>). Here, we present four examples of the activities conducted by ST-09-02 and EGIDA.

GEOSS DATA CITATION STANDARD

Geo-reference data are crucial for addressing many of the burning societal problems and to support related interdisciplinary research. Data sharing is hampered by the lack of a widely accepted method for giving credit to those who make their data freely available and for tracking the use of data throughout its life-cycle. Particularly in the scientific community, recognition and renown are important currencies. Providing means for data citation would be a strong incentive for data sharing.

A number of organizations and projects have started to address the concept of data citation (e.g., PANGAEA, NASA DAACS, USGS, NOAA National Data Centers, ESIP, the US National Academy of Sciences, and EGIDA). Several proposals for guidelines have emerged and a better understanding of the many issues at hand is evolving, but to date, no standard has been accepted. Data citation is far more complicated than citation of scientific publication because data sets differ in many aspects from standard scientific publications.

GEO is in a unique position to provide the testbed for the implementation of a draft standard with GEOSS. The S&T Committee (STC) of GEO with support of the EGIDA Project has developed a draft standard. This draft is based on guidelines developed by international groups. The key characteristics of the draft standard are summarized in Table 1. The GEO Plenary supports the testbed implementation of this draft standard.

Currently, users of the GEO-Portal are not obliged or encouraged to cite data accessed through GEOSS – if at all, citation requirements come from the individual data providers. The testbed implementation of the draft GEOSS Data Citation Standard will rectify this situation; increase the attractiveness of GEO and GEOSS for scientists by making their contributions visibly acknowledged; and help to identify issues not covered by the standard. The process of implementing the draft and iteratively improving it is led by ID-03, and coordinated with the GEO

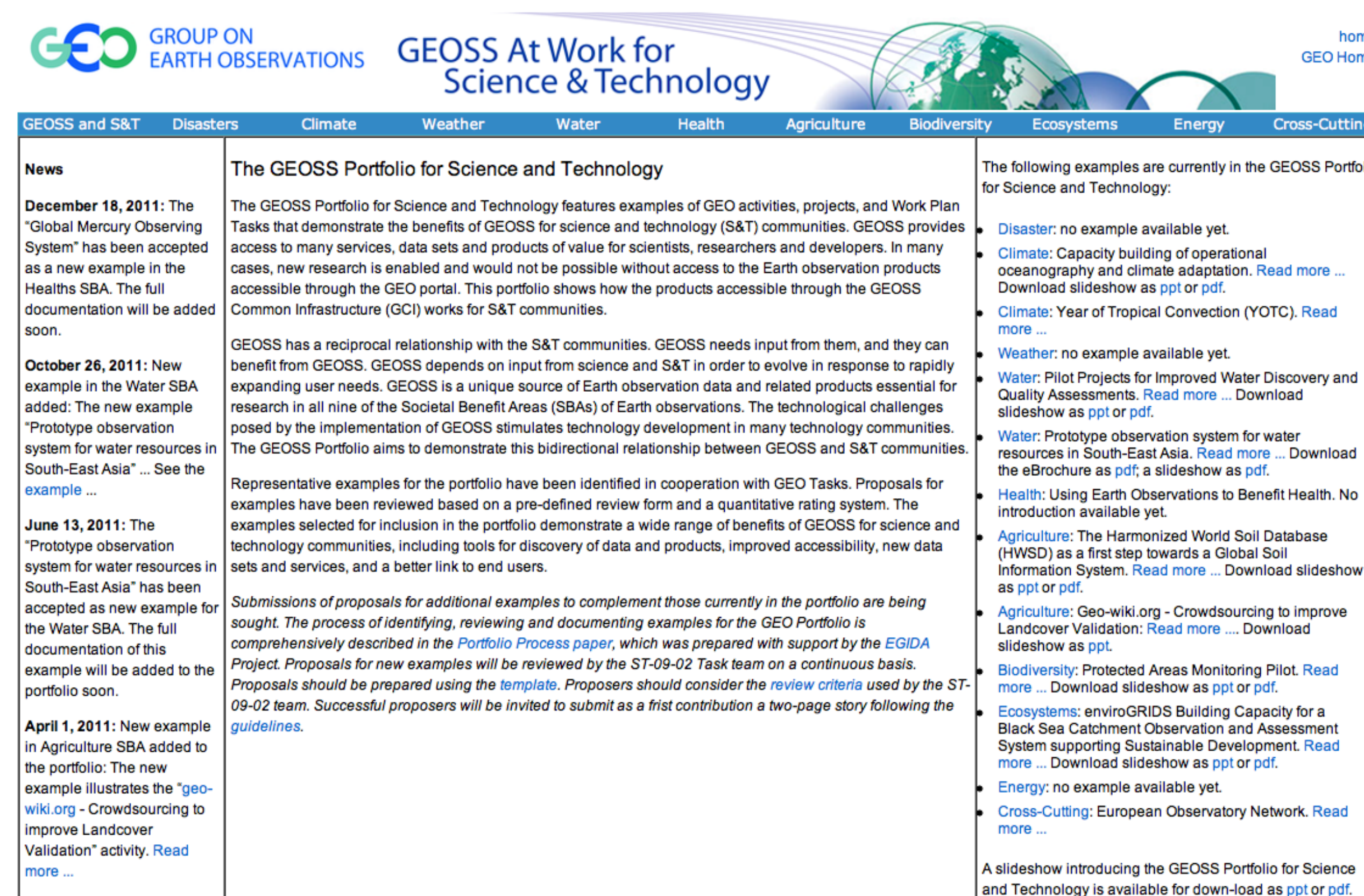


Figure 1: Home page of the GEOSS Portfolio for Science and Technology. The portfolio is accessible at http://www.geo-task.org/geoss_portfolio.

Figure 2: Home page of the Meeting Portal providing information on S&T meetings relevant for GEO and outreach activities. The page provides information on GEO activities at these meetings. Available at http://www.geo-tasks.org/meetings_sandt.

Table 1: Attributes to be included in Data Citation

Field	Description
Author or Investigator	Individual(s) whose intellectual work, such as a particular field experiment or algorithm, led to the creation of the data set.
Publication Date	For a completed data set, the publication date is simply the year of release. On-going updates to a time series do change the content of the data set, but they do not typically constitute a new version or edition of a data set. New versions typically reflect changes in sampling protocols, algorithms, quality control processes, etc. Both a new version and an update may be reflected in the publication date. The title should indicate the new version.
Title	This is the formal title of the data set. It may also include version or edition information.
Dates used	For time series, especially continually updated time series, indicate which dates of data were used. Note this is distinct from the publication date.
Editor or Compiler	An editor is the person or team who is responsible for creating a value-added and possibly quality-controlled product from the data. For minimal scientific or technical input, yet still substantial effort in compiling the product, the term compiler is more appropriate. Editors and compilers may often be responsible for a larger work that includes data sets from many individuals. Occasionally, there may be both a compiler and editor. Some products will have neither.
Publication place	This is the city, state (when necessary), and country of the publisher.
Publisher	The publisher is whoever published the data set. A publisher often has an implied responsibility for stewardship of the data set. This is usually a data center and is written immediately after the place.
Distributor or Associated Publisher	Rarely, the distributor may differ from the publisher, and in these cases, this field should be used.
Distribution Medium and Location	If there is one fixed medium, it should be listed. If data are available over the internet or through multiple digital media options it is best to include a reference to the location of the data. Often this is through a standard URL. Ideally, a persistent identifier such as a Digital Object Identifier should be used.
Access date	Because data can be dynamic and changeable in ways that are not always reflected in publication dates and versions, it is important to indicate when on-line data were accessed. It is not necessary to indicate an access date for a fixed medium like a DVD.
Data within a larger work	A particular data set may be part of a compilation, in which case it is appropriate to cite the data set somewhat like a chapter in an edited volume.

Table 2: Possible Goals of GEO Label

Community/Area	Impacts
Science and Technology	attractive incentive for involvement of S&T communities: - recognition for contributions; - enabling credits for providers (attribution); - supporting forward traceability (usage).
Users	inform users: - increase trustworthiness; - characterizing quality; - characterizing applicability; - ensuring backward traceability (data sources).
Data Sharing	promote data sharing: - signal data availability and conditions.
Providers	inform providers (and their funders): - provide information on relevance (meeting user needs); - provide information on usage; - provide user feedback on applicability

working groups in charge of developing the GEOSS Common Infrastructure; other relevant GEO components (e.g., the GEO Data Sharing Task Force); and it is aligned with the emerging international specifications concerning data citation. The experience with the testbed will be infused into the international discussion on data citation.

GEO LABEL

The availability of reliable and applicable spatial data and information is fundamental for addressing pressing problems such as food, water, and energy security; disaster risk reduction; climate change; environmental quality; pandemics; economic crises and wars; population migration; and, in a general sense, sustainability. Increasingly, tools such as Geographic Information Systems and web-based tools for viewing, accessing, and analyzing of geo-referenced information allow the combination of data sets from various sources. Improvements of interoperability, promoted particularly by GEOSS, will strengthen this trend and lead to more tools for the combinations of data from different sources. What is currently lacking is a service-oriented infrastructure helping to ensure that data quality and applicability are not compromised through modifications and combinations.

GEO has embarked on establishing a so-called GEO Label that would provide easy-to-understand, globally available information on aspects of quality, user rating, relevance, and fit-for-usage of the products and services accessible through GEOSS (with the responsibility for the concept development delegated to Work Plan Task ID-03). In designing a service-oriented architecture that could support a GEO Label, it is important to understand the impact of the goals for the label on the design of the infrastructure.

Design, concept, implementation, and success of a label depend on the goals, and these goals need to be well-defined and widely accepted. Strong labels are generally those that are unique in their field and accepted by an authoritative body in this field. A label requires time to get accepted, and once established the key characteristics normally can not be changed. An informed decision on a labeling for geo-referenced data is crucial for success. GEO is in a position to make this decision. Possible goals for the GEO Label are listed in Table 2. GEO will have to decide which of these goals to choose for the GEO Label.

GEOSS PORTFOLIO FOR SCIENCE AND TECHNOLOGY

The GEOSS S&T Portfolio features a number of compelling examples showing how GEOSS serves S&T communities in their work. The examples, which have been selected through a rigorous review of proposals submitted to ST-09-02, are accessible through a web page (Fig. 1). Currently, not all SBAs are represented through examples. The Portfolio is open for submission of additional proposals, and guidelines for the submission are available on the Portfolio web page.

GEO PRESENCE AT MAJOR CONFERENCE

GEO has made significant outreach effort to engage a larger segment of the S&T communities in GEO. A number of GEO workshops involving S&T communities have increased awareness of GEO in these communities and beyond. GEO has been present at major science meetings, organized sessions on GEOSS-related topics, convened side events, and facilitated presentations on science applications of GEOSS. To facilitate some level of coordination of these outreach activities, ST-09-02 has developed, and ID-03 is maintaining, an interactive web page for the coordination and documentation of outreach at major S&T meetings (Fig. 2). The web page also provides documentation of past GEO activities at science conferences.