

THE GEOSS USER REQUIREMENT REGISTRY: LINKING USERS OF GEOSS ACROSS DISCIPLINES AND SOCIETAL AREAS

H.-P. Plag¹, G. Foley², S. Jules-Plag³, J. Kaufman⁴, G. Ondich⁴

¹University of Nevada, Reno, Nevada Bureau of Mines and Geology and Seismological Laboratory,
Mail Stop 178, Reno, NV 89557, USA

²Environmental Protection Agency, Washington, D.C., USA

³Tiwah, Inc., Reno, NV, USA

⁴Scientific Consulting Group, Gaithersburg, MD, USA

ABSTRACT

The Group on Earth Observations (GEO) is implementing the Global Earth Observation System of Systems (GEOSS) as a user-driven service infrastructure responding to the needs of users in nine interdependent Societal Benefit Areas (SBAs) of Earth observations (EOs). The GEOSS User Requirements Registry (URR), which is part of the GEOSS Common Infrastructure (GCI), is developed as a community-based infrastructure for the collection, sharing, and analysis of user needs and EO requirements. The data model of the URR has seven core relations for User Types, Applications, Requirements, Research Needs, Infrastructure Needs, Technology Needs, and Capacity Building Needs. Semantic aspects are handled by a Lexicon, which includes a number of controlled vocabularies. With the novel concept of a Links relation, the URR captures the interconnectivity of any pair of entries in the seven core relations. Information on the societal relevance and value of a link and its implementation status supports the identification of EO priorities. The URR provides a basis for the determination of observing priorities, gap analysis, and the measuring of relevance of datasets.

Index Terms— User requirements, Earth Observations, Crowd sourcing

1. INTRODUCTION

The Global Earth Observation System of Systems (GEOSS) implemented by the Group on Earth Observations (GEO) is user-driven and responds to the needs of users in nine interdependent Societal Benefit Areas (SBAs) of Earth observations (EOs). GEOSS requires an interdisciplinary scientific approach integrating observations, research, and knowledge in these SBAs in order to enable the scientific interpretation of the collected observations and the extraction of actionable information. Using EOs to actually produce the societal benefits means getting the data to users, i.e., decision-makers, thus needing to know what they need and how they would use

it. A key infrastructure element linking users and GEOSS is the User Requirements Registry (URR), which is developed as a facility that enables a wide range of users to express their needs in terms of EOs and to understand the benefits of GEOSS for their fields. Of particular interest are users in science and technology communities, which need to be involved in both the development and the use of GEOSS. The URR is designed to capture not only the needs of these communities in terms of data and products, but also the research, infrastructure, technology, and capacity building needs that would enable new applications.

The importance of GEOSS being developed as a user-driven system of systems is emphasized in the 10-Year Implementation Plan for GEOSS [1]. In 2006, the GEO User Interface Committee (UIC) started the development of the URR with a discussion of the design, data model, and functionality of this component of the GEOSS Common Infrastructure (GCI). A prototype became available in 2010 [2]. In 2011, further progress was made in the development of the data model and the implementation of the URR [3]. The Concept of Operation (CoO) for the URR as part of the GCI was developed in 2011 [4] and finalization of the operational URR is expected at the end of 2012. By then, the URR will have menus for searching, viewing and exporting entries in the URR (View); publishing data (Publish); and analyzing networks, gaps, and relevance (Analyze). A graphical interface allows users to explore the interdependencies between entries. Users will be able to customize the URR, including the specification of a sub-domain of the URR as the work basis.

The GCI includes registries enabling users to discover, access, and use EOs and derived products and services available through GEOSS [2, 3]. The registries are the GEOSS Components and Services Registry (CSR), the GEOSS Standards and Interoperability Registry (SIR), and the Best Practices Wiki (BPW). These registries primarily focus on the providers and collect information related to the available services and products, as well as information needed to ensure interoperability between the services contributing to GEOSS. Complement-

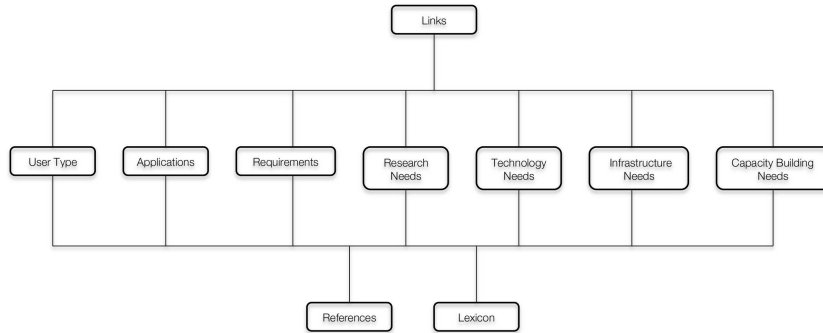


Fig. 1. Data model of the URR. The data model can be separated into three parts: (1) seven relations to capture information on user needs and applications (middle); (2) the Links relation to capture connectivity (top); (3) a number of auxiliary relations to capture relevant information used in the other parts (bottom).

ing these system-related registries, the URR is a user-related registry that allows for the collection, sharing, and analysis of user needs and EO requirements. With this, the URR enables an efficient dialog between users and providers. It is a comprehensive community-based facility to publish, view, and analyze information on societal activities depending on EOs and derived products, on the user types associated with these applications, and on a range of needs and requirements. With a novel link concept, the URR capture the interdependency between entries.

Experience with the population of the URR has been gained since 2010. The UIC carried out a review of user requirements in the nine SBAs, and this activity produced SBA-specific summary reports of observational requirements. In 2010, the reports in the Health and Disasters SBAs (available at <http://sbageotask.larc.nasa.gov/>) were used by experts to translate the requirements into the URR data model. In 2011, the use cases of risk-based management and Integrated Coastal Zone Management (ICZM) provided two examples of cross-cutting areas. The use cases were supported by the Geohazards Community of Practice (GHCP) and Coastal Zone Community of Practice (CZCP), respectively. In 2012, the Water Community of Practice added information pertaining to the Water SBA. In response to demands from users of the URR, on-line tutorials have been made available, which explain comprehensively the concepts of the URR and guide users through publishing, viewing, and analyzing the URR contents.

2. PURPOSE OF THE URR

The URR is the place in the GCI for the specification and quantification of user needs and user-related information such that this information can be utilized in analyses that give guidance to the development of a user-driven GEOSS. Development of a user-driven GEOSS requires an understanding of (1) what has priority for users; (2) which of these priorities are not met by the available datasets and products; and (3) what is the user-related relevance of existing or proposed datasets. The URR provides the tools and information to conduct three primary types of analyses corresponding to these questions.

The user-related information collected in the URR enables the three generic types of analyses: (1) prioritization of user needs and observational requirements; (2) gap analyses; (3) relevance analyses. Key for all three analyses is information on the interconnectivity between specific entries in the URR, which is captured by the entries in the Links relation (see next section). In the URR, Links entries describe the relationship between two specific entries in any of the URR relations. Links can be defined between entries in the same or in two different relations. With this versatile concept to capture interconnectivity, the data model of the URR allows for the prioritization to be carried out solely within the context of the URR. Prioritization of specific applications, observational requirements, and needs in research, infrastructure, technology and capacity building can be performed on the basis of their relevance for other entries either in the same relation or the cross links to other relations published in the URR [5].

The information published in the URR is required for one of the five approaches to gap analyses recommended by the GEO Gap Analysis Task Force [6]. Based on the interconnectivity captured by the published Link entries, it is possible to answer the question of whether an application, which depends on another application, can be carried out or whether the link is not implemented. Similarly, infrastructure, technology, capacity building and research needs, together with relevant link information, are often indicative of gaps to be filled. A gap analysis in terms of observational system gaps is based on the core question of whether a specific requirement entry can be met with existing data or products. Answering this question necessitates a search for data that meet, or exceed, the characteristics specified for the requirement entry. This search needs to be performed by other components of the GCI, specifically the Discovery and Access Broker (DAB).

In a relevance analysis, the core question is whether a specific or generic dataset or product (existing or proposed) corresponds to one or more requirements. A measure for relevance based on the number and relevance of the requirements met by a dataset is under development. It is noted here that the relevance analyses will have to be initiated from outside of the URR, e.g., through the GEOSS Web Portal.

3. THE URR DATA MODEL

The core of the data model of the URR consists of seven relations that describe the user types, applications, requirements, and various needs related to decision making and support through EOs (Fig. 1). The basic relations of the URR are:

- **Applications:** processes and activities that use EOs or derived information to produce new information, arrive at decisions, or execute decisions.
- **User Types:** generic users who are involved in applications, benefit from these, or contribute to them.
- **Requirements:** specifications of EOs or products.
- **Research Needs:** research tasks to be performed in order to enable applications that are currently not possible due to a lack of knowledge.
- **Technology Needs:** description of preferably EO requirements that cannot be met because the necessary technology to carry out the observations is not available and a description of what this technology would be.
- **Infrastructure Needs:** description of requirements that cannot be met or applications that can not take place because of the lack of infrastructure and a description of the infrastructure necessary to enable the applications.
- **Capacity Building Needs:** description of problems that cannot be solved today because of a lack of capacity and a description of the necessary capacity building that would enable applications addressing these problems.

The entries in the **Links** relation capture the interconnectivity between a source entry and a target entry in the above seven relations. This concept is a rather powerful way of capturing interconnectivity between pairs of entries. Information on the societal relevance and the implementation status of a link enable the analysis of value chains from EOs to end applications. It is planned to add the option to link URR entries to entries in the SIF, CSR, and BPW.

The **Lexicon** relation collects all terms used in the URR and manages other semantic aspects, including the definition of acronyms and abbreviations and a set of controlled vocabularies. These vocabularies include Keywords, EO Parameters, Attributes, Mediums, Units, Geographical Areas, Essential Variables, and Glossary Terms. The **References** table collects references to documents that associate entries with background information.

Entries in the seven main relations can be linked to keywords defined in the Lexicon. Requirements can be specified for EO parameters, which first need to be entered in the Lexicon and there specified as EO parameters. In some cases, it is not the parameter itself but rather a derived quantity that is needed (e.g., the concentration), and derived quantities are denoted by attributes. Some observations are only needed in a specific medium (e.g., soil moisture in soil).

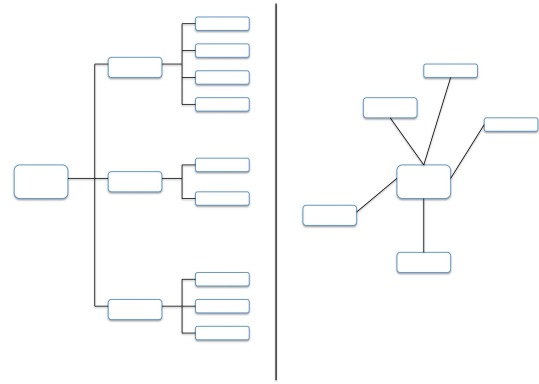


Fig. 2. Two principle approaches to the analysis of needs in a societal area. Left: the global, top-down approach resulting in a comprehensive structure for the area; right: the local, bottom-up approach starting from a local environment. The local approach can be used by any expert in a narrowly defined area of decision making. A global approach requires a broad expert knowledge of a comprehensive field.

4. METHODS FOR POPULATING THE URR

Populating the URR has two distinct aspects, one being the analysis of user needs and applications in a societal area, and the other being the collection of information to be published in the URR. There are several ways of approaching the analysis of the various needs and requirements that can be published in the URR. The approaches separate into two main classes, which we denote as global and local approaches (Fig 2). The first one is more top-down and the second one more bottom up.

In a global approach, we take a starting point at a high level, rather general entry. This could be an application, user type, requirement, etc. The tutorials use the examples of ICZM and risk-based planning to illustrate this approach (Fig. 3). A user type providing a good starting point could be “national government,” and an example for a high-level requirement would be “atmospheric methane contents.” Starting from the selected entry, we can identify all those activities and needs that related to this entry and then progress for each of them to the next level. In this approach, it makes sense to analyze first the dependencies within one group, e.g., for all applications.

In a local approach, we can take any possible entry as the starting point and analyze the local environment of this entry. As an example for this approach, the tutorials use the application Early Warning.

While a local approach lends itself to a more ad hoc start and an immediate publishing, a global approach requires considerable analysis prior to the publication of any entries in the URR. The local approach requires some effort to ensure that the local environment is linked appropriately to the existing contents in the URR and does not duplicate already existing

entries. In the global approach it is also important to use as many as possible of the existing entries and to link the determined structure with the rest of the URR at all appropriate levels.

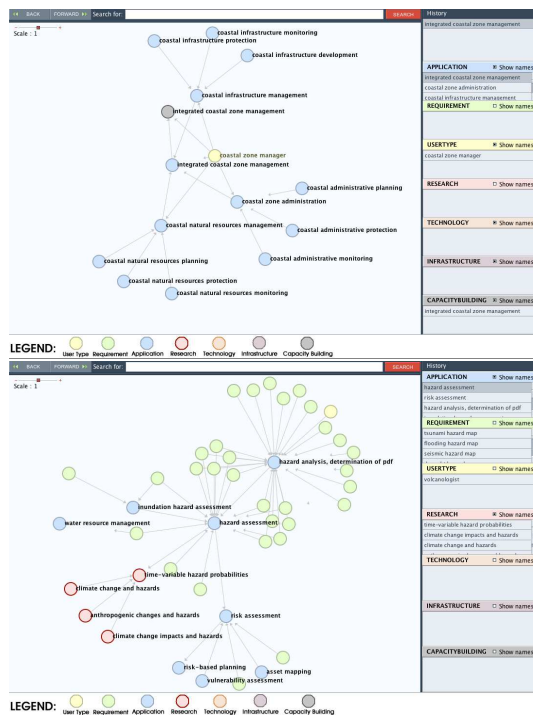


Fig. 3. Examples of graphical presentation of the interconnectivity captured in the URR. Top: center of graphics is the application ICZM; bottom: center is the application hazard assessment.

Concerning the collection of information, we have identified four different methods: (1) Experts in a societal area analyze the area using the global approach and publish the results; (2) in a crowd-sourcing approach, people publish the information relevant to their immediate environment using the local approach; (3) information published in other user requirement registries is harvested and integrated into the URR; and (4) a moderator collects information in interviews with users asking a set of agreed-upon questions. To date, methods (1) and (2) have been primarily employed in populating the URR. Method (3) is being used to harvest the WMO requirements database, and method (4) is being investigated separately, through the application of the URR infrastructure to a domestic (e.g., United States) environmental issue.

5. DISCUSSION, CONCLUSIONS AND PERSPECTIVE

The URR is a viable tool for the publishing and viewing of user-related information, and the analyzing of the links between Earth observations, applications, and users. The data

model of the URR appears to be comprehensive and flexible enough to capture all user-related aspects such that they can be compared to the system-related information in the GCI.

The URR experience concerning publishing shows that the entries in the URR display a wide range of quality, and in many cases, editing is required. The use of experts is not always a guarantee for the published information being consistent with the URR data model and of high quality. The value of the URR depends on the information published in the URR. In order to reach many potential users, the URR needs to offer an ensemble of methods for the collection of information; perhaps beyond the four described above. The URR offers a versatile model for compiling and analyzing data. The inclusion of research, infrastructure, technology and capacity building needs in the data model opens the URR for emerging uses such as the assessment of the societal relevance of a research need. By publishing and viewing problem-specific data in the URR and graphically displaying options, users may have the opportunity to exhibit and differentiate among problem-solving approaches that may not be possible without this dynamic analytical tool.

6. REFERENCES

- [1] GEO, "Global Earth Observing System of Systems GEOSS - 10-Year Implementation Plan Reference Document - Draft," Tech. Rep. GEO 1000R/ESA SP 1284, ESA Publication Division, ESTEC, Noordwijk, The Netherlands, February 2005 2005, Available at <http://earthobservations.org>.
- [2] H.-P. Plag, G. Ondich, J. Kaufman, and G. Foley, "The GEOSS User Requirement Registry - supporting a user-driven Global Earth Observation System of Systems," *Imaging Notes*, vol. 25, no. 4, pp. 28–33, 2010.
- [3] H.-P. Plag, G. Ondich, J. Kaufman, G. Foley, and F. Pignatelli, "The GEOSS User Requirement Registry: A versatile tool for the dialog between users and providers," in *Proceed. 34th Internat. Symp. Remote Sensing of the Environment, Sydney, April 10-15, 2011*, 2011.
- [4] GEO, "Concept of Operations Document User Requirements Registry (URR)," Tech. Rep., Available at http://www.scgcorp.com/urr/pdfs/URR_COO_FINAL.pdf, 2011.
- [5] H.-P. Plag, J. Pearlman, and F. Pearlman, "A measure for relevance of user needs, observational requirements, and earth observations data," Tech. Rep., GEO User Requirement Registry Development Team, 2012.
- [6] GEO Gap Analysis Task Force, "Development of a GEOSS gap analysis strategy," Doc. 13, presented to the 21st GEO Executive Committee Meeting, March 21-22, 2011, GEO Secretariat, Geneva, Switzerland, 2011.