Semantics of European Drought Regional Implementation and its Extension to Africa

William Pozzi (1), Juergen Vogt (2), Cristiano Fugazza (3), Mattia Santoro (4), Stefano Nataivi (4), Michael Piasecki (5), and Stefan Niemeyer (2)

(1) GEO AIP Water and Drought Working Group, Washington, United States (will.pozzi@gmail.com), (2) JRC/European Drought Observatory, (3) JRC, (4) IMAA CNR, (5) City College, New York (CCNY), USA

Lawrence (2010) has noted that no one simply wants to find, look at, or collect data. To the contrary, they all want something to do with data, to prepare a basis for rapid decisionmaking. The technologies of the Semantic Web (Berners-Lee, et.al.2001; Yu 2007) empower the ability to search and discover information, as well as "repackage" the information in a form more suitable for decisions. Drought monitoring, both on the European regional and the global level provide such a use for semantics.

Drought indicators identify the envelope or boundaries of drought zones, so that the course of the drought zone may be charted over time. However, drought lacks a precise and universally accepted definition, so that multiple drought indicators are necessary to more reliably document when drought occurs:drought is a "creeping" phenomenon, so that identification of the earlier stages of drought is extremely difficult (Castillo 2009; Moreira et al 2008). Hence, multiple indicators and indices have been developed in order to detect drought, the impact of drought on the various water budget components (evapotranspiration, precipitation, soil water and runoff), impact of drought upon uses of water, and remote sensing-based vegetation health indicators, designed to monitor vegetation water stress over very large tracts of land, such as the Russian Federation. In short, drought monitoring at regional and global scales calls for integrating the information of multiple disciplines, as well as semantic registration and linking of datasets to the detailed components of the water cycle. At the same time, global drought monitoring involves drought monitoring terminology among multiple languages in Europe, Africa, South America, and East Asia.

The EuroGEOSS discovery broker is a contribution to the Group on Earth Observations (GEO) and the Global Earth Observation System of Systems (GEOSS). The Discovery Augmentation Component enables semantics-aware discovery by matching the search patterns entered by the end user against a collection of multilingual, thesauri: these are controlled vocabularies providing multiple textual representations for terms and organizing them according to specificity and relatedness. As a consequence, the user's query is first related to a set of language-neutral identifiers (e.g., URIs) that represent entities in a concept graph that the user may navigate for identifying related terms that are relevant to a relief agency or drought manager's search. These data structures are hosted by the GENESIS Vocabulary Service.

These thesauri are provided in the Simple Knowledge Organizing System (SKOS) format, a lightweight ontology for expressing knowledge organization systems, harmonized in the context of GEO and EuroGEOSS project areas, by relating terms from distinct thesauri, thus allowing the user to move from one categorization to the other. The user identifies a set of terms (concept listings for drought areas), combined with drawing a box on a map, at which point the broker translates the corresponding URIs back to a customizable set of languages and executes multiple queries against the catalogs it is federating.

The General Multilingual Environmental Thesaurus (GEMET) and the INSPIRE Feature Concept Dictionary and Glossary these reference thesauri proved to be too general as to express the specificities of drought-related datasets and services, as well as details of the water cycle, that are necessary to identify whether drought has occurred, along with the severity of drought and the magnitude of its impacts. As a result, a new thematic thesaurus was established for water that could more easily be used for annotation and discovery and to bring the water thesaurus more in alignment with the drought activities, including metadata drought creation and the indicators and datasets used in common with the European Drought Observatory. The water thesaurus chosen was derived from the USA Consortium of Universities for the Advancement of Hydrologic Science, Inc (CUAHSI) water ontology, version 1. An expanded version of the CUAHSI water ontology is being translated into SKOS data structures and linked to relevant terms in the reference thesauri. The resulting system will provide concept searches, which will improve on simple keyword searches, and provide greater versatility for more diverse semi-arid drought locations, such as Africa, complimenting coverage in the more moisture-rich European areas.