

## **The Taxonomic Synonymy Definition Framework: An Implementation of the Resource Description Framework for the Exchange of Taxonomic Synonymy**

Authors: Douglas Fils<sup>1</sup>, Iain Matcham<sup>2</sup>, Cinzia Cervato<sup>1</sup>, James Crampton<sup>2</sup>

(1) Geological & Atmospheric Sciences, Iowa State University – CHRONOS, 253 Science 1 Ames, IA 50011

(2) Institute of Geological and Nuclear Science Limited, 1 Fairway Drive Avalon, Lower Hutt, 6315 New Zealand

CHRONOS in collaboration with Geological and Nuclear Sciences New Zealand (GNS) has begun development and implementation of a Taxonomic Synonymy Definition Framework (TSDF). Based on the Resource Description Format (RDF) the TSDF defines a mechanism to codify and exchange synonymy concepts in RDF using principles and concepts from the Simple Knowledge Organization System (SKOS).

TSDF utilizes as much as possible predefined community namespaces like Dublin Core, Friend of a Friend (FOAF) and others in addition to SKOS. Careful use of namespaces allows for the gleaning of data, extensibility by others and potential interaction with other synonymy encoding schema.

The goal is to provide a navigable interface to synonymy data for the researcher as well as a standard means to exchange synonymy concepts between institutions. Implementation in RDF allows for the use of faceted interfaces to RDF data such as Longwell from the MIT Simile Project as well as the rich set of RDF tools from the community. These interfaces allow for easy navigation of the various relationship paths between concepts in an RDF document. Practices from the Semantic Web Environmental Directory project, part of the Semantic Web Advanced Development (SWAD-Europe), are being used to guide the development of an architecture for a distributed community of synonymy providers. The goal is the development of a community of synonymy providers able to publish and harvest concepts from each other for use in searches or other taxonomy related research.

TSDF is a method to represent synonymy data. It is not the scope of TSDF to define a method to store or query data when the use of web services and Representative State Transfer (ReST) architectures can quite easily address the need for requesting and transporting data in TSDF format.

A query of TSDF formatted files can be done in many ways with SPARQL (recursively, SPARQL Protocol and RDF Query Language) being one of the most common methods to extract data from RDF graphs. This exchange of data is facilitated without imposing obligations on data storage or network application structures of the host institutions. Application of unique identifiers along with predefined namespaces is used for the identification of taxonomic concepts within TSDF. The use of preexisting practices provides an environment where TSDF data can be widely and easily accessed by a large community of users and providers. Details on the use of these existing practices will be given.

The presentation will cover the current state of implementation of TSDF and examples of its use to exchange data between CHRONOS and GNS as well as initial implementation of interfaces to the data collections. GNS has led the development of a set of Java based tools and API's related to TSDF and these will be outlined. CHRONOS is implementing an architecture and interface to available data using these resources that will also be detailed. An architecture using web services and ReST to arbitrate requests and transportation of TSDF encoded synonymy data is detailed in the presentation. A discussion of the application of existing architectures for the development of a distributed synonymy network will be included.