Taxon Meta-Ontology TaxMeOn
– Towards an Ontology Model for Managing Changing Scientific Names in Time

Nina Laurenne, Jouni Tuominen, Mikko Koho and Eero Hyvönen
Semantic Computing Research Group (SeCo)
Aalto University, Department of Media Technology
P.O. Box 15500, FI-00076 Aalto, Finland
University of Helsinki, Department of Computer Science

Abstract
The Semantic Web is based on machine-processable meanings, i.e. metadata and ontologies providing a new promising approach for representing and managing the scientific name system. A scientific name and the related taxonomic concept(s) form a unit in which one or the other can change. Therefore, names are unreliable identifiers for taxa. We present an ontology-based method for representing scientific names and taxonomic information that change in time. The practical goal of the system is to facilitate more accurate taxonomic metadata descriptions about names, integration of heterogeneous scientific data about organisms in different times and from different sources (publications, data bases, observations, museum collections), and to enable semantic searches, and linking in applications.

Conceptualized taxonomic information is expressed as *classes* that represent sets of individual *instances*. For example, *Cerambyx* is an *instance* of the *class* genus. *Properties* defined as relations between the *classes* tell how individuals are related to each other. For example, the genus and the author are *classes* connected by the *property* 'is described by' (e.g., *Cerambyx* is described by Linnaeus). This ongoing research introduces an ontology schema, i.e., a meta-ontology with currently ca. 20 *classes* and 70 *properties* to model information about scientific names, taxonomic concepts, authorships and taxonomic statuses. The focus is not on the characters defining taxa or on physical specimens. Terms of Darwin Core were applied when possible, but additional and more elaborate terms were needed too. The system is based on the Web Ontology Language OWL standard.

Taxonomic ranks in the model are represented as an OWL subclass hierarchy of *classes,* and individual taxa (e.g. *Cerambyx*) as *instances* of them. The taxonomic hierarchy is based on the 'part of' relation between *instances* of taxa. The 'subclass of' relation is not used here, because a taxon hierarchy semantically defines memberships of organisms in groups rather than the subclass of relation (of OWL). If subclasses were used, an *instance* of a genus would be falsely an *instance* of taxa of higher levels.

The taxonomic concepts of differing views can be mapped to each other in multiple ways. The connection between taxa can be very specific (congruent, is included/includes, overlaps with; all determined by the characters or a membership of a group) or loosely defined, which leaves the choice to add incomplete or lacking information. The taxon names and the concepts are referred to using Universal Resource Identifiers (URI) that also act as pointers to the location of the information that describe the objects on the web. A change in a taxon name or in a concept leads to the creation of a new concept and URI that is related to the former concept(s) and the time of the change. For example, when a species is shifted into another genus, a new species *instance* is created without changing the old, and the two views can be managed in time.

The ontology editor Protegé was used for constructing the *classes* and the *properties* of the model. The model is being applied to develop an ontology of nine genera of Afro-tropical eucnemid beetles. The taxonomy of the group is challenging and for instance, the genus *Pterotarsus* has gone through 22 taxonomic changes. The result will be published on the ONKI Ontology Service.

Acknowledgements
This work is part of the national FinnONTO program and is funded by the Finnish Funding Agency for Technology and Innovation (Tekes) and a consortium of 38 public organizations and companies. Fruitful co-operation with Jyrki Muona, Hans Silfverberg, Hannu Saarenmaa, and the Finnish Museum of Natural History is acknowledged.

URLs
Semantic Computing Research Group: http://www.seco.tkk.fi/
National Semantic Web Ontology Project in Finland (FinnONTO), 2003-2012: http://www.seco.tkk.fi/projects/finnonto/
Biological Ontologies and Vocabularies: http://www.seco.tkk.fi/ontologies/biology/
ONKI Ontology Service: http://www.onki.fi/