

- **GEMET:** A general multilingual thesaurus aimed to define a common language and core terminology for the environment. GEMET's data is available in SKOS (RDF/XML) format and can be either downloaded¹⁰ or accessed through the provided RESTful and XML-RPC interfaces¹¹.

3.1 SKOSification of Catalogue of Life

The extended use of taxonomies in the biodiversity domain dictates for a formal way of describing complex vocabularies and taxonomies, in compliance to the Semantic Web standards. The most popular standard for describing these types of controlled vocabularies is SKOS (Simple Knowledge Organization System) [8]. SKOS is formally described as an OWL Full Ontology, providing the basic notions and semantics needed for describing knowledge in knowledge organization systems. Its use facilitates the semantic linkage of museum objects to well-established KOS, including GEMET and Uniprot which have already been expressed in SKOS. Another system that is widely used in biological classification but is not available in SKOS format is the Catalogue of Life, which has been described above.

The current implementation of CoL provides a web-based system for browsing the taxonomy of the species, as well as services for searching, but lacks support for persistent URIs able to be referenced by external applications, and RDF representation of its data. Towards this end, we have worked on a method of exposing the taxonomy of CoL to RDF, and more specifically SKOS, using the annual checklist, which is a downloadable package containing the relational database of the CoL.

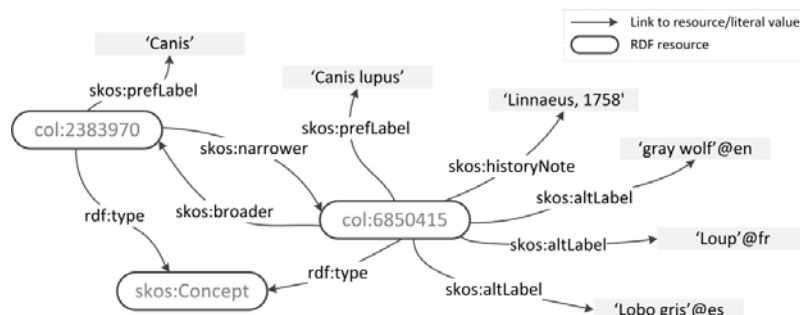


Fig. 2. An example of the Catalogue of Life SKOSified data in the form of a graph

For the conversion of the CoL dataset to SKOS we used the D2R Server [1], which allows the publishing of relational databases in RDF format. The features of the SKOS model that we employed are: (a) the class *Concept*, and (b) the properties *broader*, *narrower*, *prefLabel* and *altLabel*. The first step was the representation of all the taxonomy nodes as *Concepts*. The scientific name of each node was transformed into a *prefLabel*, and the common names into *altLabels*. Finally, the hierarchy of the

¹⁰ <http://www.eionet.europa.eu/gemet/rdf>

¹¹ <http://taskman.eionet.europa.eu/projects/zope/wiki/GEMETWebServiceAPI>

