

The Birds of the World Ontology AVIO

Jouni Tuominen, Nina Laurenne, Mikko Koho, and Eero Hyvönen

Semantic Computing Research Group (SeCo)
Aalto University and University of Helsinki
<http://www.seco.tkk.fi>, firstname.lastname@aalto.fi

Abstract. We present an ontology for managing the scientific and common names of birds. The ontology is based on the TaxMeOn meta-ontology model for biological names. The ontology is in use as an ontology service and it has been applied in a bird watching system.

Biologists refer to organisms by using scientific names although most people use common (vernacular) names. Modeling and managing scientific and common names is not trivial [2, 3] because names change over time. For example, a species may be shifted into another genus, which is reflected in the scientific name.

We present “The Birds of the World Ontology AVIO”¹ encompassing all bird taxa of the world, including their scientific names and common names in English, Finnish, and Swedish. The ontology is an application of our earlier work, the meta-ontology model TaxMeOn for biological names and classifications [3]. The aim of the AVIO ontology is to provide a reference to the species names of the birds, e.g., for indexing bird related information.

The AVIO ontology is available as open data and as an operational ontology service in the Finnish Ontology Library Service ONKI [4]. The ontology has been used in an online mobile birding system BirdWatch² [1].

AVIO Ontology The AVIO Ontology is based on the Finnish names of the birds of the world published by BirdLife Finland³, containing the common names for all the species of the world including their higher taxa—9,740 species, 1,227 genera, and 194 families. The English names and the scientific names are included, too. A CSV formatted species list was provided by BirdLife, and was converted into an RDF-based ontology conforming to the TaxMeOn schema. The classification of taxa was completed with additional higher level taxa.

The CSV formatted species list contained a scientific name and its common name equivalents per line. The hierarchy between the scientific names was expressed using hierarchically numbered identifiers for the names (e.g., 061.0449. *Sterna* and 061.0449.1664 *Sterna paradisaea*). In the RDF conversion, every scientific and common name was given a URI and the relations between them

¹ This work is part of the FinnONTO project (<http://www.seco.tkk.fi/projects/finnonto/>) funded mainly by Tekes.

² <http://www.demo.seco.tkk.fi/birdwatch/>

³ <http://www.birdlife.fi/>

were made explicit: 1) the hierarchical relations between the scientific names, and 2) the equivalence relations between the scientific and common names.

The scientific names are modelled as instances of the TaxMeOn classes *TaxonInChecklist* and an appropriate taxonomic rank (e.g., *Species*). The common names are expressed as instances of the class *VernacularName*. Giving separate URIs for scientific and common names supports temporal tracking of the names as the individual names can be assigned statuses with time stamps.

The resulting ontology is published in the ONKI service⁴. AVIO is available also in SKOS⁵ to support SKOS-based applications. Here the preferred and alternative scientific and common names of a taxon are represented simply as literals that cannot have further properties, unlike in TaxMeOn. This version also includes Swedish common bird names obtained from BirdLife.

Ontology Service and a Use Case ONKI provides ready-made services for accessing the AVIO ontology. End-users can browse and search the ontology, e.g., to get an overview of the classification of birds or to find a scientific name for a taxon that they know only by the common name. The ONKI selector widget can be integrated into legacy CMS systems to provide an autocomplete and URI fetching features. The ontology is published for machines as Linked Data, as a SPARQL endpoint, and there are several APIs to use, e.g., for query expansion.

AVIO ontology is used as a conceptual hub of the BirdWatch system based on linked data. This system helps citizen ornithologists in species identification, based on the ontological knowledge and the statistics of earlier observations. The URIs are used for linking bird data from different services, such as Wikipedia, the multilingual NatureGate⁶, and 1.25 million bird observations in a SPARQL endpoint based on the GBIF data⁷.

References

1. Hyvönen, E., Alonen, M., Koho, M., Tuominen, J.: BirdWatch—supporting citizen scientists for better linked data quality for biodiversity management. In: Workshop on Semantics for Biodiversity (S4BIODIV), ESWC 2013, Montpellier, France. CEUR Workshop Proceedings (2013)
2. Page, R.: Taxonomic names, metadata, and the semantic web. *Biodiversity Informatics* 3, 1–15 (2006)
3. Tuominen, J., Laurence, N., Hyvönen, E.: Biological names and taxonomies on the semantic web – managing the change in scientific conception. In: Proceedings of the ESWC 2011, Heraklion, Greece. pp. 255–269. Springer–Verlag (2011)
4. Viljanen, K., Tuominen, J., Hyvönen, E.: Ontology libraries for production use: The Finnish ontology library service ONKI. In: Proceedings of the ESWC 2009, Heraklion, Greece. pp. 781–795. Springer–Verlag (2009)

⁴ online browsing: <http://onki.fi/en/browser/overview/linnut>
download: <http://tinyurl.com/linnut-download>

⁵ <http://onki.fi/en/browser/overview/avio>

⁶ <http://www.naturegate.fi>

⁷ <http://www.gbif.org/>