

Abstract – Regular Paper – iEvoBio Conference 2010

Title: Biological taxonomy and ontology development: scope and limitations

Authors: Nico M. Franz¹ and David Thau²

¹ Presenter; Department of Biology, Call Box 9000, University of Puerto Rico, Mayagüez, PR 00681-9000 U.S.A., nico.franz@upr.edu

² Department of Computer Science, University of California at Davis, Kemper Hall, One Shields Avenue, Davis, CA 9561 U.S.A., thau@learningsite.com

Abstract. Biologically oriented ontologies are widely regarded as essential for achieving semantic integration of a vast range of biological data. Having originated in the biomedical sector, the development of ontologies is branching out into other biological disciplines including behavior, ecology, evolution, morphology, and now also taxonomy and phylogenetics (cf. Schulz *et al.* 2008. *Bioinformatics* 24: i313-i321). Nevertheless, actual ontological implementations of taxonomies in popular biodiversity information services are rather simplistic, and consequently are unable to support complex taxonomic searches or permit reasoning among concepts that pertain to alternative taxonomic perspectives. Here we critically examine to what extent biological taxonomies are suitable for ontology-based representation. We show that there are at least three domains – nomenclature, classification and phylogeny – for which alternative ontologies and inference services may fruitfully be developed. Focusing on the latter two domains, we then illustrate through a series of practical examples why taxonomies pose a series of potentially unique challenges that limit the scope of ontology development. These include: the inherently non-static nature of taxonomies, their dependence on a natural hierarchy which exists (in some sense) independently of taxonomic practice, the evolvability of taxonomic properties, the hybrid class-/individual-like nature of taxonomic entities, the critical yet idiosyncratic semantic properties of Linnean ranks, the indelibility of nomenclatural and taxonomic legacies, the insufficient exploration of the tree of life, and the need to perform increasingly complex concept alignment services both under a static and dynamic perspective of taxonomic data. We conclude that the most fruitful path to integrating ontologies into taxonomic and phylogenetic research should focus on developing more powerful algorithms and tools to produce such concept alignments. We also point out that the success of ontological applications in systematic research will depend largely on the willingness of the expert community to annotate its products in such a way that concept alignments become more feasible.