



Principles for a names-based cyberinfrastructure to serve all of biology*

DAVID J. PATTERSON^{1,3}, SARAH FAULWETTER² & ALEXEY SHIPUNOV¹

¹Biodiversity Informatics Group, Encyclopedia of Life, Marine Biological Laboratory, 7 MBL Street, Woods Hole, MA 02543, USA

²Institute of Marine Biology and Genetics, Hellenic Centre for Marine Research, 71003 Heraklion, Crete, Greece

³E-mail: dpatterson@eol.org

* In: Minelli, A., Bonato, L. & Fusco, G. (eds) *Updating the Linnaean Heritage: Names as Tools for Thinking about Animals and Plants*. Zootaxa, 1950, 1–163.

Table of contents

Abstract	152
Introduction	154
The principles.....	155
Inclusive	155
Comprehensive.....	155
Taxonomically intelligent reconciliation	156
Taxonomically intelligent disambiguation.....	157
Concept-capable.....	157
Hierarchical structure	158
Phylogenetic structure.....	158
Distributed organization.....	159
Interoperability.....	159
Up-to-date	159
Participation	161
Authoritative	161
Scalable	162
Conclusions	162
Acknowledgements	162
References	162

Abstract

The rapidly growing amount of biological data on the internet and the increasing need for large-scale analyses mandate improvements to the management of taxon-centric information. This information, traditionally managed by taxonomists, is now transforming into a web-based infrastructure. The complexity and narrative quality of the biological sciences require an information management framework that is sensitive to the scale, richness, character, and heterogeneity of the discipline. Given that the names of organisms offer us a nearly universal system for indexing biological data objects, a names-based cyberinfrastructure has the capacity to index the totality of available biological information and to aggregate taxon-centric data over a broad scale. In order to serve its role, this infrastructure should incorporate thirteen principles that are proposed here.

Key words: Biodiversity informatics, Taxonomic intelligence, Encyclopedia of Life, Data management, Taxonomy