The Large Taxon Collider: Building the Atlas of Living Australia's taxonomic backbone

Doug Palmer‡, Peggy Newman§

‡ Atlas of Living Australia, Canberra, Australia
§ Atlas of Living Australia, Melbourne, Australia

Corresponding author: Doug Palmer (doug.palmer@csiro.au), Peggy Newman (peggy.newman@csiro.au)

Abstract

One of the most important components of the Atlas of Living Australia (ALA) is our taxonomic backbone, comprised of an index of species names with a suite of web services for matching names and delivering information for our species pages.

The index is built by merging lists from multiple Australian authoritative taxonomic sources into a single taxonomic tree. Where the primary data sources are incomplete, we attempt to pad out missing genera and species with alternative sources, for example, using Catalog of Life for some fungi branches, and classification of kingdoms: Viruses, Chromista, Protozoa and Bacteria.

When the ALA ingests occurrence records or species checklists, we attempt to match the supplied names against the index using our name matching service. All going well, the service returns the authoritative persistent identifiers for the matched taxon concept and all of the parent taxon concepts. The name matching service will attempt fuzzy and higher rank matches in an attempt to ensure that every record in the ALA finds a home in the taxonomic tree.

The merging and matching algorithms are available as open-source code.

This presentation will give a brief overview of the construction of our names index, the technical and taxonomic challenges the ALA has faced in building and delivering the service, and our development plans for the future.

Keywords

species lists, taxonomy

Presenting author

Doug Palmer

Presented at

TDWG 2022