The Reptile Database: Curating the biodiversity literature without funding

Peter Uetz ‡

‡ Virginia Commonwealth University, Richmond, United States of America

Corresponding author: Peter Uetz (peter@uetz.us)

Abstract

The Reptile Database (RDB) curates the literature and taxonomy for about 14,000 species and subspecies of reptiles (Uetz et al. 2021). Together with a few other databases, the RDB curates the literature for about 70,000 species of fish, amphibians, reptiles, birds and mammals.

While it acts as a current name list for extant reptile taxa, including synonymies, it also collects images (currently ~18,000, representing half of all species), type information, diagnoses and descriptions, and a bibliography of 62,000 references, most of which are linked to online sources.

The database is also extensively cross-referenced to citizen science projects (iNaturalist), the NCBI taxonomy, the IUCN Red List, and several others, and serves as data provider (for reptiles) for the Catalogue of Life.

A major challenge for the Reptile database is the consistent curation of the literature, which requires the addition of about 2000 papers a year, including about 200 new species descriptions and numerous taxonomic changes. For instance, during the past five years, almost 1000 species changed their names, in addition to the ~900 species that were newly described, i.e., almost 20% of all reptile species were described or changed their name within just a half decade!

While the database can keep track of name changes, it remains a largely unsolved problem of how these name changes can or should be translated into related databases such as the National Center for Biotechnology Information (NCBI), which keeps track of the literature independently (but exchanges data with the RDB). Some sites use the web services of the RDB to update their taxonomy, such as Calphotos or iNaturalist, but many do not or have not been able to implement automated name tracking.

The RDB also works with the Global Assessment of Reptile Distributions (GARD Initiative) to keep track of range changes. After GARD published a collection of ~10,000 range maps

©

.

for reptiles in 2017, more than half of these maps have changed in area size by more than 5% since the initial release.

The database has developed several avenues for streamlining and optimizing curation of the literature, e.g., (semi-) automated requests for publications, species descriptions, and photos from authors, but the process is far from fully automated.

Questions remain: how can taxonomic databases develop, share, and exchange better tools for curation? Can we standardize data collection and processing? How can we automatically exchange data with other data sources? How can we optimize the process of scientific publication to streamline databasing and automated information extraction?

Presenting author

Peter Uetz

Presented at

TDWG 2021

Conflicts of interest

none

References

- Uetz P, Koo M, Aguilar R, Brings E, Catenazzi A, Chang AT, Chaitanya R, Freed P, Gross J, Hammermann M, Hošek J, Lambert M, Sergi Z, Spencer CL, Summers K, Tarvin R, Vredenburg VT, Wake DB, et al. (2021) A Quarter Century of Reptile and Amphibian Databases. Herpetologica Review 52 (2): 246-255. [In English]. URL: https://www.researchgate.net/publication/352462027_A_Quarter_Century_of_Reptile_and_Amphibian_Databases