

Extinct Taxa in an Extant World: Working towards better fossil taxonomic representation

Holly Little[‡], Christina Byrd[§], Talia Karim[|], Erica Krimmel[¶], Ben Norton[#]

[‡] Smithsonian National Museum of Natural History, Washington, DC, United States of America

[§] Harvard Museum of Comparative Zoology, Cambridge, Massachusetts, United States of America

[|] University of Colorado, Boulder, Colorado, United States of America

[¶] iDigBio, Florida State University, Tallahassee, Florida, United States of America

[#] North Carolina Museum of Natural Sciences, Raleigh, North Carolina, United States of America

Corresponding author: Holly Little (littleh@si.edu)

Abstract

Paleontology collections face many similar challenges mobilizing taxonomic data as extant collections disciplines: from complex nomenclatural histories to incomplete taxonomic authorities to questionable specimen identifications. These challenges exist within the information management of an individual institution and are amplified at the level of data aggregation, where a minor misspelling of authorship or a historic name can completely erase the discoverability of a specimen. In addition to those challenges shared with extant taxonomy, fossil taxonomy must also address uncertainty in far more detail and at many levels. It may be impossible to assign a name into the traditional Linnaean hierarchy or to provide classification at higher ranks due to uncertainty in our understanding of how the specimen is related to other known taxa. Moreover, some branches of paleontology prefer not to use Linnaean ranks at all, making it even more difficult to map concepts and/or clade names to terms in the Linnaean-based Darwin Core standard (DwC, Darwin Core Task Group 2009). Incomplete preservation of fossil specimens also makes it difficult or impossible to identify many specimens to the extant gold standard rank of species. While these specimens are often recorded as precisely as possible (e.g., to a secondary Linnaean rank like subfamily) in local collection management systems, the data are obscured when shared with global aggregators since these ranks do not exist in the DwC framework.

The paleontology collections community has been facilitating efforts to address these challenges for many years in an attempt to improve the quality of our data and the treatment of taxonomic information for fossil specimens by aggregators. Most recently, the [Paleo Data Working Group](#) has been focusing on the practical matters of digitizing, mobilizing, and promoting the use of specimen data from our collections, and further assessing possible community-sourced solutions that could increase the discoverability of our specimens from a taxonomic entry point.

In this talk, we share our explorations into the past and current state of fossil taxonomy in the global landscape (including notable gaps) and its use in paleontology collections, and we also report on possible paths forward. Understanding some of the historical issues with fossil taxonomy in the global landscape has led to a better understanding of root differences between extant and fossil taxonomic data models. The Paleo Data Working Group (Krimmel et al. 2021) has reviewed fossil taxonomic data as it is published through aggregators like the Global Biodiversity Information Facility ([GBIF](#)), and identified concrete issues in order to classify systematic pain points. We have discussed existing sources of taxonomic data and the agency of collections professionals to take an active role in working toward new solutions for enhancing and evolving these systems. Opportunities for intentional action have been identified across a spectrum of local to global efforts, and from individual capacity to collaboration across stakeholder groups, connecting with those who can more directly address a particular gap or combine expertise. On the individual scale, we are promoting the adoption of best practices for recording taxonomic data in collections information systems and subsequent mapping to DwC. Through coordination with stakeholder groups, we are hoping to enable paleontologists to provide taxonomic authority information to [Catalogue of Life](#) for certain fossil groups in new ways. Working within this space, we have tested new systems for collaborative taxonomic data curation and produced updated community data guidelines for use in paleontology collections, to improve the consistency and quality of mobilized data.

Accurately representing extinct taxa in an extant world is a moving target, affected by constantly evolving data standards and technical capacity. For instance, GBIF is moving to an operational model where Catalogue of Life provides the taxonomic backbone, but fossil names are problematic in Catalogue of Life because they cannot always integrate into the single tree of life perspective. Seeking solutions that are inclusive of fossil taxonomy will enable our global corpus of biodiversity data to extend back into deep time and promote new opportunities for synergy and cooperation.

Keywords

fossil taxonomy, paleobiology, Catalogue of Life, GBIF backbone taxonomy

Presenting author

Holly Little

Presented at

TDWG 2022

Conflicts of interest

References

- Darwin Core Task Group (2009) Darwin Core. Biodiversity Information Standards (TDWG). URL: <http://www.tdwg.org/standards/450>
- Krimmel E, Karim T, Little H, Walker LJ, Burkhalter R, Byrd C, Millhouse A, Utrup J (2021) The Paleo Data Working Group: A model for developing and sustaining a community of practice. Biodiversity Information Science and Standards 5: e74370. <https://doi.org/10.3897/biss.5.74370>