

The iDigBio US Collections List: Now powered by GBIF

Cat Chapman ‡

‡ University of Florida, Gainesville, United States of America

Corresponding author: Cat Chapman (cchapman@flmnh.ufl.edu)

Abstract

[iDigBio](#) (Integrated Digitized Biocollections), the US national biodiversity data aggregator, publishes a [list of US Collections](#) that is intended to be a comprehensive list of natural science collections in the United States of America. This list aims to provide access to information and metadata about natural science collections in the United States, including but not limited to, collections descriptions, contact information, taxonomic scope of collections, and links to existing recordsets within iDigBio (if applicable). Previously, this list was maintained as a JSON (JavaScript Object Notation) endpoint via GitHub, with updates maintained manually, requiring substantial human involvement and reliance on third party services (e.g., TravisCI) to publish new collections entries or updates to existing collections metadata. This was a time consuming and fragile process; if GitHub or TravisCI became unavailable or nonfunctional for any reason, updates to the US Collections List could not be published.

In 2020, the iDigBio US Collections List [was successfully merged](#) with [GRSciColl](#), the Registry of Scientific Collections at the [Global Biodiversity Information Facility \(GBIF\)](#). GRSciColl and the US Collections List fundamentally share the same goal: enhancing access to information about natural science collections, associated digitized recordsets, and personnel involved with these collections. The US Collections List is now maintained directly on GRSciColl by GBIF and iDigBio staff, and the US Collections List hosted on [iDigBio.org](#) is now populated via the [GRSciColl Application Programming Interface \(API\)](#). This merger has resulted in a more streamlined experience for both those maintaining the list and users of the list; changes submitted to US entries on GRSciColl now appear instantaneously on the US Collections List at iDigBio. Engaging the broader community is fundamental for data integrity; GRSciColl has implemented functionality for transparent requests for metadata changes (e.g., change in contact information) from GRSciColl users. These changes are evaluated by GRSciColl maintainers before publishing; approved changes are visible immediately. Furthermore, GRSciColl has connectivity with [Index Herbariorum](#), a ledger of herbarium collections around the world. Notably, one feature of GRSciColl is the concept of a "Master Record", where an entry on GRSciColl can have an "authoritative record" from a source external to GRSciColl; an entry on Index Herbariorum

is one example of such a Master Record. This connectivity with authoritative sources such as Index Herbariorum increases community cohesion and accuracy of data provided to and by GRSciColl and, ultimately, the US Collections List.

We hope that this unified global index of natural science collections will continue to enhance access to information about biodiversity collections and the people and data involved.

Keywords

metadata, service, index, community, global, information

Presenting author

Cat Chapman

Presented at

TDWG 2022

Acknowledgements

This work is supported by the US National Science Foundation as part of cooperative agreement DBI-2027654. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

The author would also like to acknowledge: Marie Grosjean and the rest of the GBIF Cyberinfrastructure team for all their help and insight; the iDigBio Cyberinfrastructure team for their work in implementing the necessary changes on the iDigBio backend.

Conflicts of interest