

Conference Abstract

An Introduction to the Global Registry of Scientific Collections (GRSciColl)

Melissa Jean-Yi Liu^{‡,§}, Lily Shrestha[‡], Marie Grosjean[¶], Morten Høfft[¶], Marcos Lopez Gonzalez[¶], Tim Robertson[¶]

‡ Asia Regional Support Team, Global Biodiversity Information Facility, Taipei, Taiwan

§ Taiwan Biodiversity Information Facility, Biodiversity Research Center, Academia sinica, Taipei, Taiwan

‡ Asia Regional Support Team, Global Biodiversity Information Facility, Kathmandu, Nepal

¶ Global Biodiversity Information Facility, Copenhagen, Denmark

Corresponding author: Marie Grosjean (mgrosjean@gbif.org)

Received: 08 Aug 2024 | Published: 12 Aug 2024

Citation: Liu MJ-Y, Shrestha L, Grosjean M, Høfft M, Gonzalez M, Robertson T (2024) An Introduction to the Global Registry of Scientific Collections (GRSciColl). Biodiversity Information Science and Standards 8: e134257. <https://doi.org/10.3897/biss.8.134257>

Abstract

The Global Registry of Scientific Collections ([GRSciColl](#)) is a comprehensive community-curated repository of scientific collections, their host institutions, and associated staff members. GRSciColl was adopted by the Global Biodiversity Information Facility ([GBIF](#)) in 2018, after being developed and hosted at the Smithsonian Institution.

GRSciColl houses data on various institutional or personal scientific collections such as natural history collections, archeological collections, and cell culture collections. Users can access information about the collection location, content, accessibility status, coverage (e.g., geographic and taxonomic), and contacts. Additionally, it serves as a tool for enhancing data interoperability by including institution and collection codes and identifiers, as well as external identifiers like Wikidata Q numbers whenever feasible. The digitized specimens published on [GBIF.org](#) are linked with GRSciColl institutions and collections based on these codes and identifiers. These specimens are then aggregated to generate metrics and dashboards, which display customized multiple tables and visualization, regardless of how they were published on GBIF.

Most of the data in GRSciColl is sourced from [Index Herbariorum](#), the Integrated Digitized Biocollections ([iDigBio](#)), and national registries. The information can also be improved by direct contributions from a community of editors or the suggestions made by the users.

As of 19 April 2024, GRSciColl provided information on 7,678 scientific collections and 7,906 institutions. However, GRSciColl includes numerous legacy entries, with many institutions lacking associated collections. This discrepancy arises partly from misregistrations, where institution entries should be collections. Additionally, some national registries or surveys provided institution-level information that did not align with the GRSciColl schema for collections. Furthermore, some records pertain to historical institutions that no longer maintain collections. The most up-to-date entries are discoverable and accessible via the website: <https://scientific-collections.gbif.org> and the [GRSciColl Application Programming Interface](#) (API).

In this presentation, we will show you how to navigate the GRSciColl website and API to help you find relevant institutions, collections, and specimen records and look up codes and identifiers. We will cover any new functionality developed in the context of the [GRSciColl 2023-2024 road map](#), which aims to support richer, structured descriptions as well as the ability to upload an inventory of species.

Keywords

collection repository, data curation, data mobilization

Presenting author

Melissa Jean-Yi Liu

Presented at

SPNHC-TDWG 2024

Conflicts of interest

The authors have declared that no competing interests exist.