Best Practices for Research Data Sharing

Daniel Noesgaard | GBIF Secretariat
CHORUS Forum, 27 March 2024

- **Datasets**: 103,397
- **Organizational Participants**: 43
- **Country Participants**: 63
- **Hosted portals**: 20
- **Peer-review papers using data**: 10,257
- **Average records downloaded per month (2023)**: 173.8 billion
- **Species occurrence records**: 2,688,299,408
- **Publishers**: 2,252
Enabling FAIR data sharing and citations in GBIF

**Findable**

GBIF assigns DOIs through DataCite to every dataset available through GBIF, but also to queries/downloads of records across multiple datasets.

GBIF offers the ability to mints DOIs for derived datasets of GBIF data not accessed directly through GBIF web services (e.g. cloud services, 3rd party libraries, etc.)

**Accessible**

Datasets and their metadata is available through web pages, REST API (JSON) and also through DataCite

Several community-developed software libraries available (R, Python, etc.)

Metadata and landing pages remain even if datasets are deleted

**Interoperable**

Data can be downloaded in several formats including simple row-oriented CSV format, but also column-oriented Parquet format

 Appropriately licenced datasets are also directly available for analysis on several cloud computing environments, e.g. the MS Planetary Computer, Amazon AWS and Google Cloud

**Reusable**

GBIF-mediated data comes from a variety of sources, but all data shared through the network are standardized using the Darwin Core Standard (~200 fields)

DwC is a community-developed biodiversity data standard, governed and ratified by Biodiversity Information Standards (aka TDWG)
Peer-reviewed publications using GBIF-mediated data

DOI-based citations of GBIF-mediated data
Collectively develop tools, processes and incentives

- Publishers, editors, reviewers, etc.
- Improve data citation guidelines
- Automatatic workflows to flag “bad” citations
- Enforce DOIs for ALL citations of GBIF-mediated data
Thank you!

Daniel Noesgaard | dnoesgaard@gbif.org
@daniel@social.data.coop

#CiteTheDOI

EVERY time!