Domain/Commercial Context

Mark Leggott, Executive Director **|** RDC NDSF Summit **|** Sep 22/17

Let's connect: mark.leggott@rdc-drc.ca **|** @mleggott

RDC 🍁 DRC

Researcher Interest

> Types of services of interest
> - Creation & Storage Services
> - Computational & Analysis Services
> - Discovery Services
> - Identifier Services
> - Dissemination Services
> - Preservation Services
> - Support and Training Services

Interest Focus

> Depositing data as appropriate to respond to publisher/funder/institutional/personal requirements

> Goal for all is to increase:
  - Discoverability
  - Reproducibility
  - Reusability
  - Citability

> Also matches community of practice interest

Publisher – Scientific Data

> "Data should be submitted to discipline-specific, community-recognized repositories where possible, or to generalist repositories if no suitable community resource is available."

> "[repositories] have been evaluated to ensure that they meet our requirements for data access, preservation and stability."

> Data must be included with submission, or not considered further.

Recommended Data Repositories

View data repositories

- Biological sciences:
  - nucleic acid sequence; protein sequence; molecular & supramolecular structure; neuroscience; omics; taxonomy & species diversity; mathematical & modelling resources; cytometry & immunology; organism-focused resources

- Health sciences
- Chemistry & chemical biology
- Earth and environmental sciences
- Physics, astrophysics & astronomy
- Social sciences
- Generalist repositories
- Other repositories

Recommended **Data Repositories**

Taxonomy & species diversity

Integrated Taxonomic Information System (ITIS)	view FAIRsharing entry
KNB: The Knowledge Network for Biocomplexity	view FAIRsharing entry
NCBI Taxonomy *	view FAIRsharing entry
Global Biodiversity Information Facility (GBIF)	view FAIRsharing entry
Morphobank.org	view FAIRsharing entry

GBIF | Global Biodiversity Information Facility

Free and open access to biodiversity data

OCCURRENCES   SPECIES   DATASETS   PUBLISHERS   RESOURCES

Search

WHAT IS GBIF?   ABOUT GBIF CANADA

British soldier lichen *Cladonia cristatella* by Bob Rodrigez via iNaturalist. Photo licensed under CC BY-NC 4.0.

Occurrence records
849,734,069

Datasets
36,397

Publishers
1,102

Species
Learn more about the number of species covered by data in GBIF.org.

Fernandes becomes second consecutive Brazilian to win GBIF Young Researchers Awa...
21 September 2017

Using social media to complement traditional sources of biodiversity data
12 September 2017

Escribano becomes first Spaniard to win GBIF Young Researchers Award
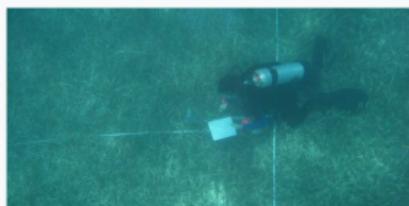21 September 2017

Job opportunities for software developers
14 September 2017

Top tips for the new GBIF.org
Updates on new features (last updated 13 Sep 2017)

GBIF Content Mobilization Priorities 2017
Current priorities for guiding investment and effort to digitize and mobilize biodiversity data

Introduction to sampling-event data
What is sampling-event data and why is it important?

The

Doze
toget
acce

Global Biodiversity Information Facility

Abbreviation: GBIF

General Information

GBIF, the Global Biodiversity Information Facility, is an international open data infrastructure, funded by governments. It is not a data repository. Rather, GBIF indexes thousands of datasets shared freely by hundreds of institutions worldwide making it the biggest biodiversity database on the Internet. GBIF.org makes the data discoverable and citable, assigning each download a DOI and storing it for an extended period of time. More than 1 peer-reviewed research publication citing GBIF as a source of data is published every day, in studies spanning the impacts of climate change, the spread of pests and diseases, priority areas for conservation and food security. To share your data with GBIF follow this quick guide: http://www.gbif.org/publishing-data/quick-guide To deposit/host your data, it is recommended you use one of the trusted IPT data hosting centres (DHC) listed in FAIRsharing.org. The DHC will provide you with an account on their IPT, which will allow you to manage your own datasets and publish them through GBIF.org.

Homepage http://www.gbif.org

Developed in Worldwide

Created in 2001

Taxonomic range

`All`

Scope and data types

`Biodiversity` `Bioinformatics Tools And Formats` `Life Science` `Taxonomic Classification`

Collected/Recommended By

PLOS · SCIENTIFIC DATA · GigaScience · BioMed Central The Open Access Publisher · bioCADDIE · GBIF Global Biodiversity Information Facility

This record is maintained by gbifs `ORCID` and kbraak `ORCID`

Associated recommendations

PLOS

Scientific Data

GigaScience

BioMed Central

In Collections

Data Citation Implementation

Global Biodiversity Information Facility (GBIF) international open data infrastructure

Support

General

blog	http://gbif.blogspot.dk
contact form	http://www.gbif.org/contact/contact-us
support email	info@gbif.org
support email	helpdesk@gbif.org
mailing list	http://lists.gbif.org/mailman/listinfo/a ...
online documentation	http://www.gbif.org/disclaimer/datashari ...
twitter	@GBIF

Additional Information

Contact	GBIF Secretariat

Related Standards

Reporting Guidelines

No guidelines defined

Terminology Artifacts

Darwin Core

Models and Formats

TDWG Access Protocol for Information Retrieval
Access to Biological Collection Data
Ecological Metadata Language

Related Databases

Scratchpads
Brazilian Flora 2020
Reflora Virtual Herbarium
Atlas of Living Australia IPT - GBIF Australia Repository
GBIF Spain IPT - GBIF Spain Repository
Canadensys IPT - GBIF Canadensys Repository
SiB Colombia IPT - GBIF Colombia Repository
Belgian BIF IPT - GBIF Belgium Repository
GBIF Mauritania IPT - GBIF Belgium
Antabif IPT - AntOBIS IPT - GBIF Belgium
Biofresh IPT - GBIF Belgium
Danish Biodiversity Information Facility IPT - GBIF Denmark
Finnish Biodiversity Information Facility Digitarium IPT - GBIF Finland
GBIF France IPT - GBIF France
GBIF Benin IPT - GBIF France
GBIF Guinea IPT - GBIF France
Madagascar Biodiversity Information Facility IPT - GBIF France
GBIF Togo IPT - GBIF France
Norway IPT - GBIF Norway
Portugal IPT - GBIF Portugal
South African National Biodiversity Institute IPT - GBIF South Africa

Tools

Integrated Publishing Toolkit	http://www.gbif.org/ipt	v2.3.4

Access / Retrieve Data

Conditions of Use

Applies to: Data use

Creative Commons Attribution-NonCommercial (CC-BY-NC) 4.0 International
Creative Commons Public Domain Dedication License (CC0) 1.0
Creative Commons Attribution (CC-BY) 4.0 International

Applies to: Database software

Open-source (Apache 2.0 license)

Data Access

Literature Tracking	http://gbif.org/mendeley
Search Occurrence	http://www.gbif.org/occurrence
How To Publish	http://www.gbif.org/publishing-data/summary
Search Dataset	http://www.gbif.org/dataset
Search Species	http://www.gbif.org/species
Browse By Country	http://www.gbif.org/country

REST Web Services

GBIF API http://www.gbif.org/developer/summary	http://api.gbif.org/v1/

Taiwan Biodiversity Information Facility IPT - GBIF Taiwan
Ocean Biogeographic Information System USA IPT - USGS
Nonindigenous Aquatic Species NAS IPT - USGS
Biodiversity Information Serving Our Nation IPT - USGS
VertNet IPT
iDigBio IPT
Ocean Biogeographic Information System IPT
Australian Antarctic Data Centre

Implementing Policies

Scientific Data's Recommended Data Repositories
Scientific Data's Data Policies
GigaScience - Minimum Standards of Reporting Checklist
PLOS Recommended Data Repositories
BioMed Central - Minimum Standards of Reporting Checklist

Credit

Record Maintainer

This record is maintained by gbifs ORCID and kbraak ORCID

Maintains

Global Biodiversity Information Facility (GBIF), Copenhagen, Denmark

Publications

No Publications Added.

Generalist Repositories

Repository Name	Information on fees/costs	Size limits	Integrated with *Scientific Data's* manuscript submission system	Re3data / BioSharing entry
Dryad Digital Repository	$120 USD for first 20 GB, and $50 USD for each additional 10 GB	None stated	**Yes** ✔	view FAIRsharing entry
figshare	100 GB free per *Scientific Data* manuscript. Additional fees apply for larger datasets	1 TB per dataset	**Yes** ✔ - To qualify for the 100 GB of free storage, data must be uploaded to figshare via our submission system. Download instructions.	view FAIRsharing entry
Harvard Dataverse	Contact repository for datasets over 1 TB	2.5 GB per file, 10 GB per dataset	No	view re3data entry
Open Science Framework	None stated	5 GB per file, multiple files can be uploaded	No	view FAIRsharing entry
Zenodo	Cost information	50 GB per dataset	No	view re3data entry

Other Repositories

> "If none of the discipline-specific repositories listed above are suitable, *Scientific Data* is glad to support the use of other repositories, if they are able to mint DataCite DOIs for hosted data, and share data under open terms of use (for example the CC0 waiver)."

> "institutional, national and project-specific data repositories."

> Repositories must support confidential peer review, or Generalist option used for workflow.

Other Interests

> Researcher identifies most closely with domain approach and community of practice

> Institutional/Regional/National repositories support:
  - Promotion/discovery of scholars/outputs
  - Preservation of outputs
  - Policy (e.g. impact of funding)
  - "Other Repositories" role

RDC 🍁 DRC

Research Data Canada – Données de Recherche Canada

rdc-drc.ca | @rdc_drc

Contact me:

mark.leggott@rdc-drc.ca | @mleggott