



Deutsche  
Forschungsgemeinschaft



JACOBS  
UNIVERSITY

# GFBio: FDM für Biologie und Umweltwissenschaften

Ivaylo Kostadinov, Ph.D.

Jacobs University Bremen | Max-Planck-Institut für Marine Mikrobiologie

✉ [i.kostadinov@jacobs-university.de](mailto:i.kostadinov@jacobs-university.de)

🐦 @tigroumaniac

MPDL FDM Workshop | 2018-04-20 | München



🐦 @GFBio\_Project

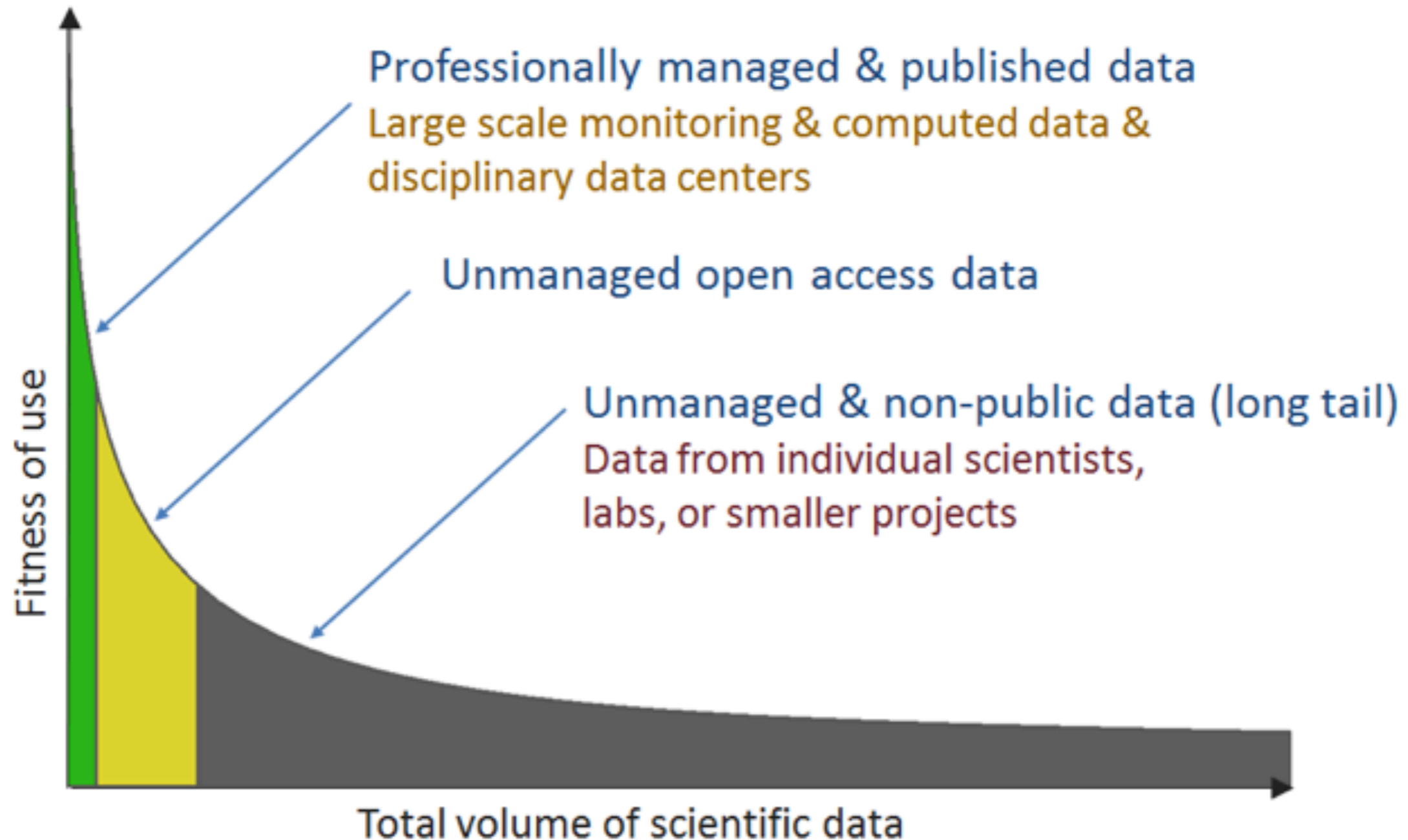
# dark data



**[...] data that has never been published or otherwise made available to the rest of the scientific community.**

B. P. Heidorn Libr. Trends 57, 280–299; 2008

# Data challenges



courtesy: Michael Diepenbroek

- Heterogeneous data
- Complex metadata standards
- Incentives for good RDM still unclear



[www.gfbio.org](http://www.gfbio.org)



A **sustainable, service-oriented** data infrastructure to facilitate **data sharing** and **data-intensive** science in biology and environmental research.

Funded by:



# Interdisciplinary team



- About ▾
- Data ▾
- Training ▾
- Support ▾
- News
- Contact

➔ Sign In

## CONSORTIUM

GFBio's multidisciplinary consortium is represented by experts from a variety of institutions ranging from natural history collections, libraries, bioinformatics to environmental data archives



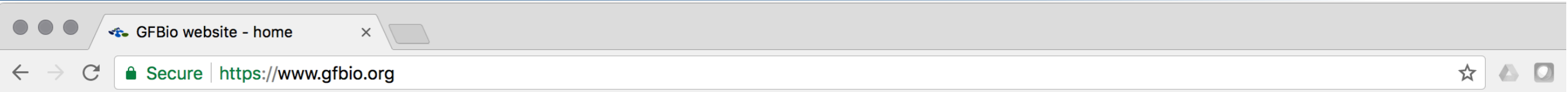
data centers and publishers

infrastructure partners

user-representatives



<http://www.gfbio.org/training/data-lifecycle>



About ▾

Data ▾

Training ▾

Support ▾

News ▾

Contact

GFBio e.V.

➔ Sign In

## SERVICES

The Key Features of our Work



### PLAN

Prepare a custom Data Management Plan (DMP).



### SUBMIT

Submit your data to GFBio.



### SEARCH

Search the GFBio data pool.



### VISUALIZE & ANALYZE

Dynamically integrate, analyze and visualize GFBio datasets.



### PUBLISH

Make your data citable.



### TRAIN

Train your data management skills.



### ARCHIVE

Deposit data and specimens in dedicated long-term archives.



### ANNOTATE & CONNECT

Use the GFBio Terminology Service to describe your data and share terminologies with other researchers.



## Submit Your Data to a Public Repository

Transfer your data from your private research domain to the [GFBio data centers](#) for long-term archival and publication. Our curation experts will find the best solution for storing your data within the GFBio consortium and making it [FAIR](#) (findable, accessible, interoperable and re-usable).

[Start a data submission](#)

### Submit Molecular Sequence Data

Molecular sequence data are submitted to the [European Nucleotide Archive](#), any accompanying environmental data are archived in [PANGAEA](#). We will also help you apply the [MixS standard](#) to your meta-data.

[Start a Molecular Submission](#)

### Data Center Recommendation Tool

Use our Data Center Recommendation Tool (DCRT) to find out which of the [GFBio data centers](#) fits best for your research data. You can also read their detailed information and even contact each data center directly.

[Start the Data Center Recommendation Tool](#)

Browser tabs: Data Center Recommendation x Data Center Recommendation x

Address bar: Secure | https://www.gfbio.org/data/dcrt

Navigation: About ▾ Data ▾ Training ▾ Support ▾ News ▾ Contact GFBio e.V. Welcome Ivaylo! Sign Out

## Data Center Recommendation Tool

Do you want to submit physical objects along with your data?

Yes  
 No

Is your object dead or alive?

Alive  
 Dead






Is your object taxon-based?

Yes  
 No

Which kind of material would you deliver?

Botanical objects

### Data Center Recommendation

	Botanic Garden and Botanical Museum Berlin, Freie Universität Berlin	<a href="#">Contact</a> <a href="#">Submission</a> <a href="#">Details</a>
	Senckenberg Gesellschaft für Naturforschung - Leibniz Institute, Frankfurt	<a href="#">Contact</a> <a href="#">Submission</a> <a href="#">Details</a>
	Staatliche naturwissenschaftliche Sammlungen Bayerns - SNSB IT Center, München	<a href="#">Contact</a> <a href="#">Submission</a> <a href="#">Details</a>
	State Museum of Natural History Stuttgart	<a href="#">Contact</a> <a href="#">Submission</a> <a href="#">Details</a>
	German Federation for Biological Data (GFBio)	<a href="#">Contact</a>

Do you need support in selecting a suitable data center or do you have further questions concerning data management?  
Please get in contact with us:

## BGBM – Botanic Garden and Botanical Museum Berlin, Freie Universität Berlin

The [Botanic Garden and Botanical Museum Berlin](#), Freie Universität Berlin with its [Research Group Biodiversity Informatics](#) is a centre of biodiversity research in Europe, housing extensive scientific collections of herbarium specimens (about 3.5 million), one of the world's largest living plants collections, a DNA Bank as well as the most complete botanical library in Germany.

Contact in the GFBio context: Anton Güntsch, David Fichtmüller and Maren Gleisberg

[Contact the BGBM GFBio team](#)

[BGBM – extended profile](#)

### Service Description

Data archiving for research projects is focusing on

**Type 1a** Botanical specimen data; (Physical) DNA Storage (under discussion at BGBM); Referenced multimedia objects

**Type 1b** Botanical observational data; Referenced multimedia objects

**Type 2** Botanical systematics and monographic works

**Type 4** RAW data (data sets and/or data packages) only if well documented and in formats and structures appropriate for long-term archiving, without further data management required

A preference will be given on data which fall under the geographic and taxonomic research foci of the BGBM ([BGBM - extended profile](#), [BGBM: Research](#)). The data archiving and publication includes management processes with [JACQ](#), [reBIND](#) -workflow and the [EDIT Platform for Cybertaxonomy](#) as well as the data quality service platform and transformation and import services provided by BGBM.

### Service Levels

Data Set	Data Package	Data Management	Research Objects
✓	✓	✓	✓

### Data Submission Formats

**Data** Preferably via Excel templates provided by BGBM and/or standardized formats of any kind (e.g. ABCD or DwC-A files); Spreadsheets (CSV, excel-files, image files); Export files from external EDIT platform installations

**Metadata** EML (GFBio submission package leaflet), ABCD, DarwinCore, DublinCore, SDD

## Data Center Recommendation



Botanic Garden and Botanical Museum Berlin, Freie Universität Berlin

[Contact](#)

[Submission](#)

[Details](#)



Senckenberg Gesellschaft für Naturforschung - Leibniz Institute, Frankfurt

[Contact](#)

[Submission](#)

[Details](#)



Staatliche naturwissenschaftliche Sammlungen Bayerns - SNSB IT Center, München

[Contact](#)

[Submission](#)

[Details](#)



State Museum of Natural History Stuttgart

[Contact](#)

[Submission](#)

[Details](#)

Do you need support in selecting a suitable data center or do you have further questions concerning data management?  
Please get in contact with us:



German Federation for Biological Data (GFBio)

[Contact](#)



**DCRT Contact Request for BGBM**
✕

**Name**

**E-Mail**

**Message**

Your Message

**BGBM – Botanic Garden and Botanical Museum Berlin, Freie Universität Berlin**

The [Botanic Garden and Botanical Museum Berlin](#), Freie Universität Berlin with its [Research Group Biodiversity Informatics](#) is a centre of biodiversity research in Europe, housing extensive scientific collections of herbarium specimens (about 3.5 million), one of the world's largest living plants collections, a DNA Bank as well as the most complete botanical library in Germany.

Contact in the GFBio context: Anton Güntsch, David Fichtmüller and Maren Gleisberg

**Service Description**

Data archiving for research projects is focusing on

**Type 1a** Botanical specimen data; (Physical) DNA Storage (under discussion at BGBM); Referenced multimedia objects

**Type 1b** Botanical observational data; Referenced multimedia objects

**Type 2** Botanical systematics and monographic works

**Type 4** RAW data (data sets and/or data packages) only if well documented and in formats and structures appropriate for long-term archiving, without further data management required

A preference will be given on data which fall under the geographic and taxonomic research foci of the BGBM ([BGBM - extended profile](#), [BGBM: Research](#)). The data archiving and publication includes management processes with [JACQ](#), [reBiND](#) -workflow and the [EDIT Platform for Cybertaxonomy](#) as well as the data quality service platform and transformation and import services provided by BGBM.

**Service Levels**

Data Set	Data Package	Data Management	Research Objects
✓	✓	✓	✓

**Data Submission Formats**

**Data** Preferably via Excel templates provided by BGBM and/or standardized formats of any kind (e.g. ABCD or DwC-A files); Spreadsheets (CSV, excel-files, image files); Export files from external EDIT platform installations

**Metadata** EML (GFBio submission package leaflet), ABCD, DarwinCore, DublinCore, SDD

### Data Center Recommendation

	Botanic Garden and Botanical Museum Berlin, Freie Universität Berlin	<input type="button" value="Contact"/> <input type="button" value="Submission"/> <input type="button" value="Details"/>
	Senckenberg Gesellschaft für Naturforschung - Leibniz Institute, Frankfurt	<input type="button" value="Contact"/> <input type="button" value="Submission"/> <input type="button" value="Details"/>
	Staatliche naturwissenschaftliche Sammlungen Bayerns - SNSB IT Center, München	<input type="button" value="Contact"/> <input type="button" value="Submission"/> <input type="button" value="Details"/>
	State Museum of Natural History Stuttgart	<input type="button" value="Contact"/> <input type="button" value="Submission"/> <input type="button" value="Details"/>
<p>Do you need support in selecting a suitable data center or do you have further questions concerning data management? Please get in contact with us:</p>		
	German Federation for Biological Data (GFBio)	<input type="button" value="Contact"/>

**DCRT Contact Request for BGBM**
✕

**Name**

**E-Mail**

**Message**

Your Message

**BGBM – Botanic Garden and Botanical Museum Berlin, Freie Universität Berlin**

The [Botanic Garden and Botanical Museum Berlin](#), Freie Universität Berlin with its [Research Group Biodiversity Informatics](#) is a centre of biodiversity research in Europe, housing extensive scientific collections of herbarium specimens (about 3.5 million), one of the world's largest living plants collections, a DNA Bank as well as the most complete botanical library in Germany.

Contact in the GFBio context: Anton Güntsch, David Fichtmüller and Maren Gleisberg

**Service Description**

Data archiving for research projects is focusing on

- Type 1a** Botanical specimen data; (Physical) DNA Storage (under discussion at BGBM); Referenced multimedia objects
- Type 1b** Botanical observational data; Referenced multimedia objects
- Type 2** Botanical systematics and monographic works
- Type 4** RAW data (data sets and/or data packages) only if well documented and in formats and structures appropriate for long-term archiving, without further data management required

A preference will be given on data which fall under the geographic and taxonomic research foci of the BGBM ([BGBM - extended profile](#), [BGBM: Research](#)). The data archiving and publication includes management processes with [JACQ](#), [reBiND](#)-workflow and the [EDIT Platform for Cybertaxonomy](#) as well as the data quality service platform and transformation and import services provided by BGBM.

**Service Levels**

Data Set	Data Package	Data Management	Research Objects
✓	✓	✓	✓

**Data Submission Formats**

**Data** Preferably via Excel templates provided by BGBM and/or standardized formats of any kind (e.g. ABCD or Dwc-A files); Spreadsheets (CSV, excel-files, image files); Export files from external EDIT platform installations

**Metadata** EML (GFBio submission package leaflet), ABCD, DarwinCore, DublinCore, SDD

**Data Center Recommendation**

Botanic Garden and Botanical Museum Berlin, Freie Universität Berlin

Submission

Dataset

Internal

Submission ID: 0000015f-da95-918e-0000-015fda95918e

Recommended Data SNSB, SMNS, BGBM, SGN

Center:

DCRT Input: {"alive": "false", "material": "Botanical objects", "taxon": "true", "physical": "true"}

Do you need support in selecting a suitable data center or do you have further questions concerning data management? Please get in contact with us:

German Federation for Biological Data (GFBio)



## Submit Your Data to a Public Repository

Transfer your data from your private research domain to the [GFBio data centers](#) for long-term archival and publication. Our curation experts will find the best solution for storing your data within the GFBio consortium and making it [FAIR](#) (findable, accessible, interoperable and re-usable).

[Start a data submission](#)

### Submit Molecular Sequence Data

Molecular sequence data are submitted to the [European Nucleotide Archive](#), any accompanying environmental data are archived in [PANGAEA](#). We will also help you apply the [MixS standard](#) to your meta-data.

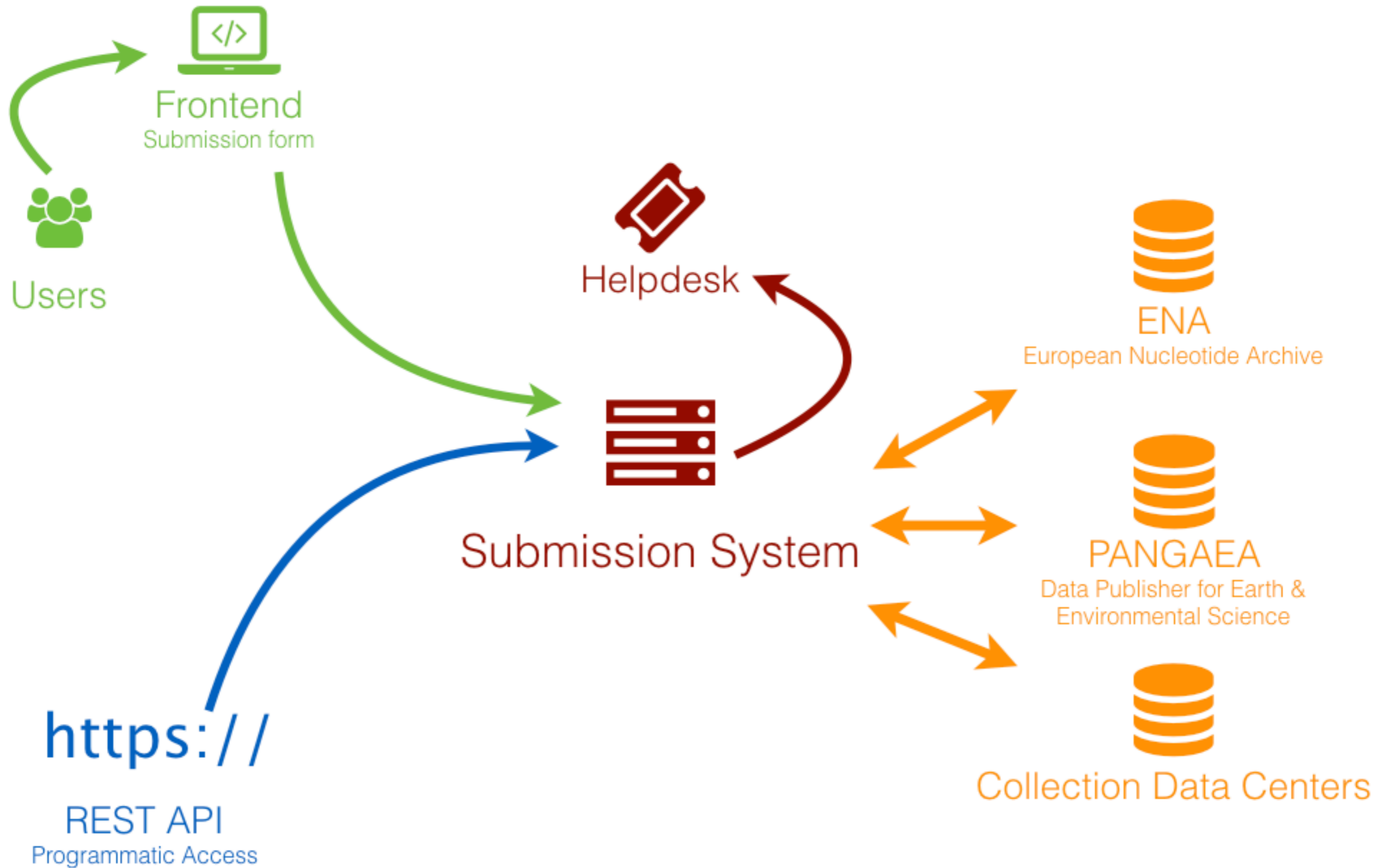
[Start a Molecular Submission](#)

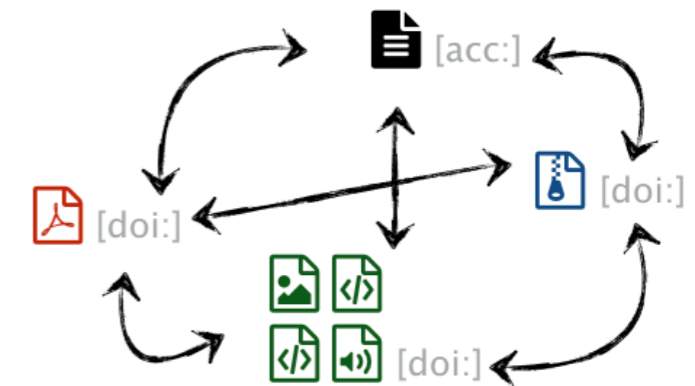
### Data Center Recommendation Tool

Use our Data Center Recommendation Tool (DCRT) to find out which of the [GFBio data centers](#) fits best for your research data. You can also read their detailed information and even contact each data center directly.

[Start the Data Center Recommendation Tool](#)

# Data Submission



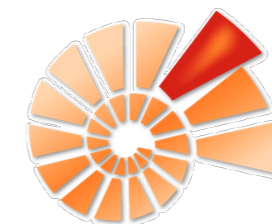


Added value at a glance:

- Single-point of contact - data is distributed to data centers and interlinked
- Support for metadata standardization - ABCD, MIxS, ENVO experts
- Manual and automated (API) operation
- Integration with local RDM systems
- Legal Framework Support



<http://bexis2.uni-jena.de/>



<http://diversityworkbench.net>





About ▾

Data ▾

Training ▾

Support ▾

News ▾

Contact

GFBio e.V.

➔ Sign In

## SERVICES

The Key Features of our Work



### PLAN

Prepare a custom Data Management Plan (DMP).



### SUBMIT

Submit your data to GFBio.



### SEARCH

Search the GFBio data pool.



### VISUALIZE & ANALYZE

Dynamically integrate, analyze and visualize GFBio datasets.



### PUBLISH

Make your data citable.



### TRAIN

Train your data management skills.



### ARCHIVE

Deposit data and specimens in dedicated long-term archives.



### ANNOTATE & CONNECT

Use the GFBio Terminology Service to describe your data and share terminologies with other researchers.

VAT: Visualization, Analysis and ... Search - GFBio website

Secure | <https://www.gfbio.org/data/search?filter=%5B%7B%22term%3A%7B%22licenseFacet%3A%22License%3A%20CC%20BY%204.0%22%7D%7D%5D>

gfbio About ▾ Data ▾ Training ▾ Support ▾ News ▾ Contact GFBio e.V. Sign In

Filter Results: [clear filters](#) [reset search](#)

License: CC BY 4.0 ✕

Geographical Region

- [Europe](#)(4515729)
- [Germany](#)(4266653)
- [Starnberg Nord \(MTB 7934\)](#)(59952)
- [Africa](#)(11749)

Search:

Show 10 entries per page

Showing 1 to 10 of 4,606,862 entries

[Previous](#) [Next](#)

[Bolbitius vitellinus \(Persoon: Fries\) Fries, a human observation record of the "IBF Monitoring of Fungi" dataset \[ID: 407030 / 472011\]](#)

**Data Center:** Data Center SNSB

**Summary:** [http://www.diversitymobile.net/wiki/IBFfungicoll\\_About](http://www.diversitymobile.net/wiki/IBFfungicoll_About)

**Parameters:** Locality; Date; Latitude; Longitude

**Dataset Citation:** Universität Regensburg, IBF Monitoring of Fungi

**License/Rights:** Permission is given for free use of data in the search result for non-commercial purposes. Any commercial use must be granted by prior written agreement.  
License: CC BY 4.0

[Data Description](#)

---

[Hebeloma crustuliniforme \(Bulliard: Fries\) Quélet non ss. Ri., Bres., a human observation record of the "IBF Monitoring of Fungi" dataset \[ID: 407030 / 472010\]](#)

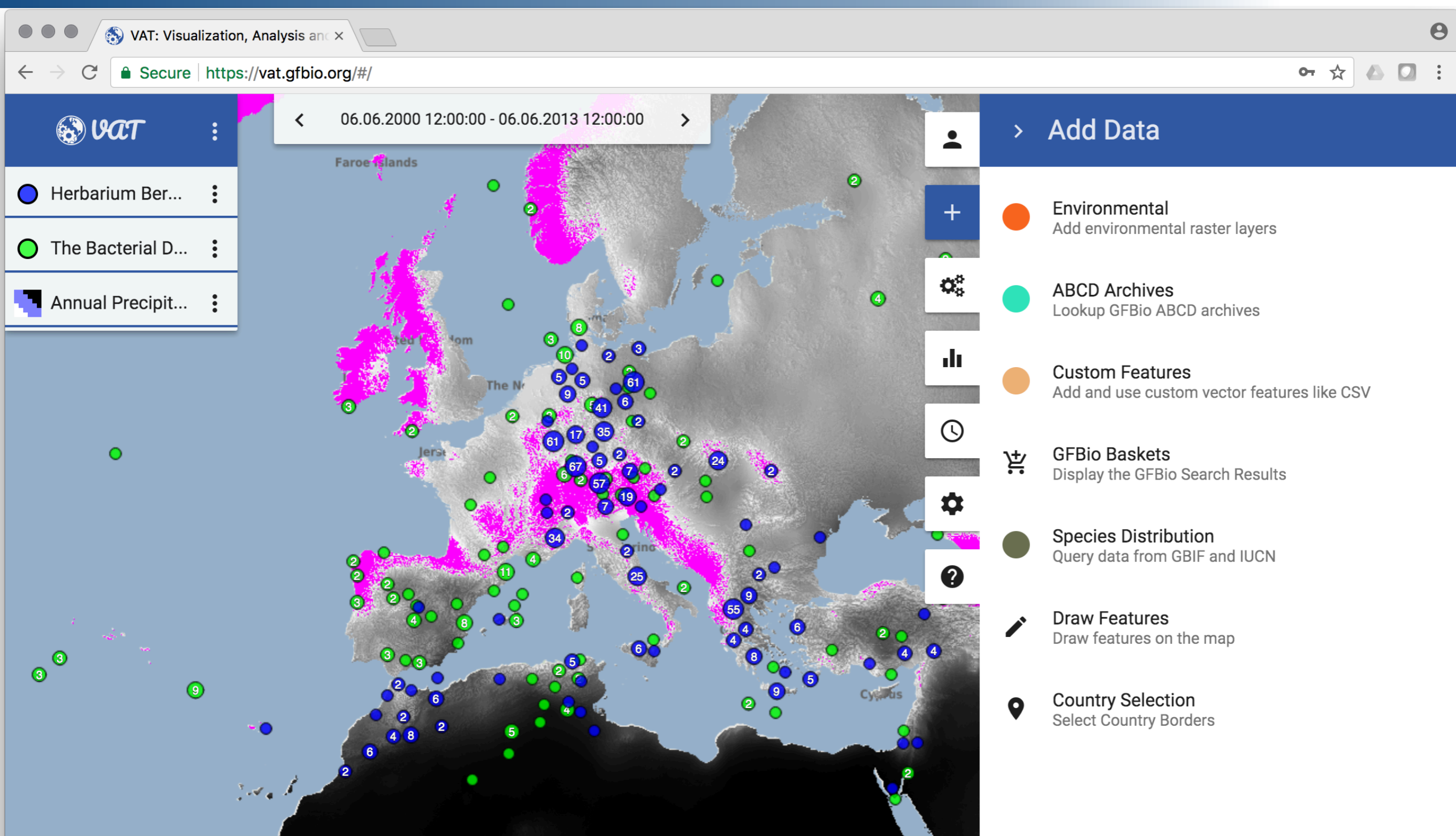
**Data Center:** Data Center SNSB

**Summary:** [http://www.diversitymobile.net/wiki/IBFfungicoll\\_About](http://www.diversitymobile.net/wiki/IBFfungicoll_About)

**Parameters:** Locality; Date; Latitude; Longitude

**Dataset Citation:** Universität Regensburg, IBF Monitoring of Fungi

# Visualization & Analysis



VAT: Visualization, Analysis and Tools

Secure | <https://vat.gfbio.org/#/>

06.06.2000 12:00:00 - 06.06.2013 12:00:00

**VAT**

- Herbarium Ber...
- The Bacterial D...
- Annual Precipit...

**Add Data**

- Environmental  
Add environmental raster layers
- ABCD Archives  
Lookup GFBio ABCD archives
- Custom Features  
Add and use custom vector features like CSV
- GFBio Baskets  
Display the GFBio Search Results
- Species Distribution  
Query data from GBIF and IUCN
- Draw Features  
Draw features on the map
- Country Selection  
Select Country Borders

# Terminology Service



https://terminologies.gfbio.org



[Browse](#) [Search](#) [API](#) [Widgets](#) [About](#)



## Explore

Browse and search for terminologies that can be interesting for your research.



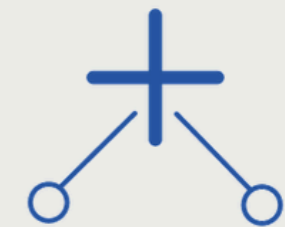
## Access

Use information from terminologies programmatically to provide semantically enriched applications.



## Consume

Retrieve and store Information from terminologies in your local information system.



## Contribute

Store or connect your terminologies to the TS and get access to all provided services automatically.



About ▾

Data ▾

Training ▾

Support ▾

News ▾

Contact

GFBio e.V.

➔ Sign In

## SERVICES

The Key Features of our Work



### PLAN

Prepare a custom Data Management Plan (DMP).



### SUBMIT

Submit your data to GFBio.



### SEARCH

Search the GFBio data pool.



### VISUALIZE & ANALYZE

Dynamically integrate, analyze and visualize GFBio datasets.



### PUBLISH

Make your data citable.



### TRAIN

Train your data management skills.



### ARCHIVE

Deposit data and specimens in dedicated long-term archives.

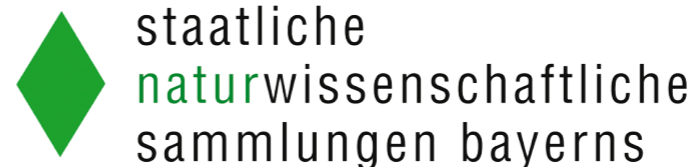


### ANNOTATE & CONNECT

Use the GFBio Terminology Service to describe your data and share terminologies with other researchers.



gfbio e.V. [anno 2016]



# Thank You!

 [www.gfbio.org](http://www.gfbio.org)

 [info@gfbio.org](mailto:info@gfbio.org)

 [@GFBio\\_Project](https://twitter.com/GFBio_Project)

Questions?