

Managing Research Data^{❄️}

Ivaylo Kostadinov, Ph.D.

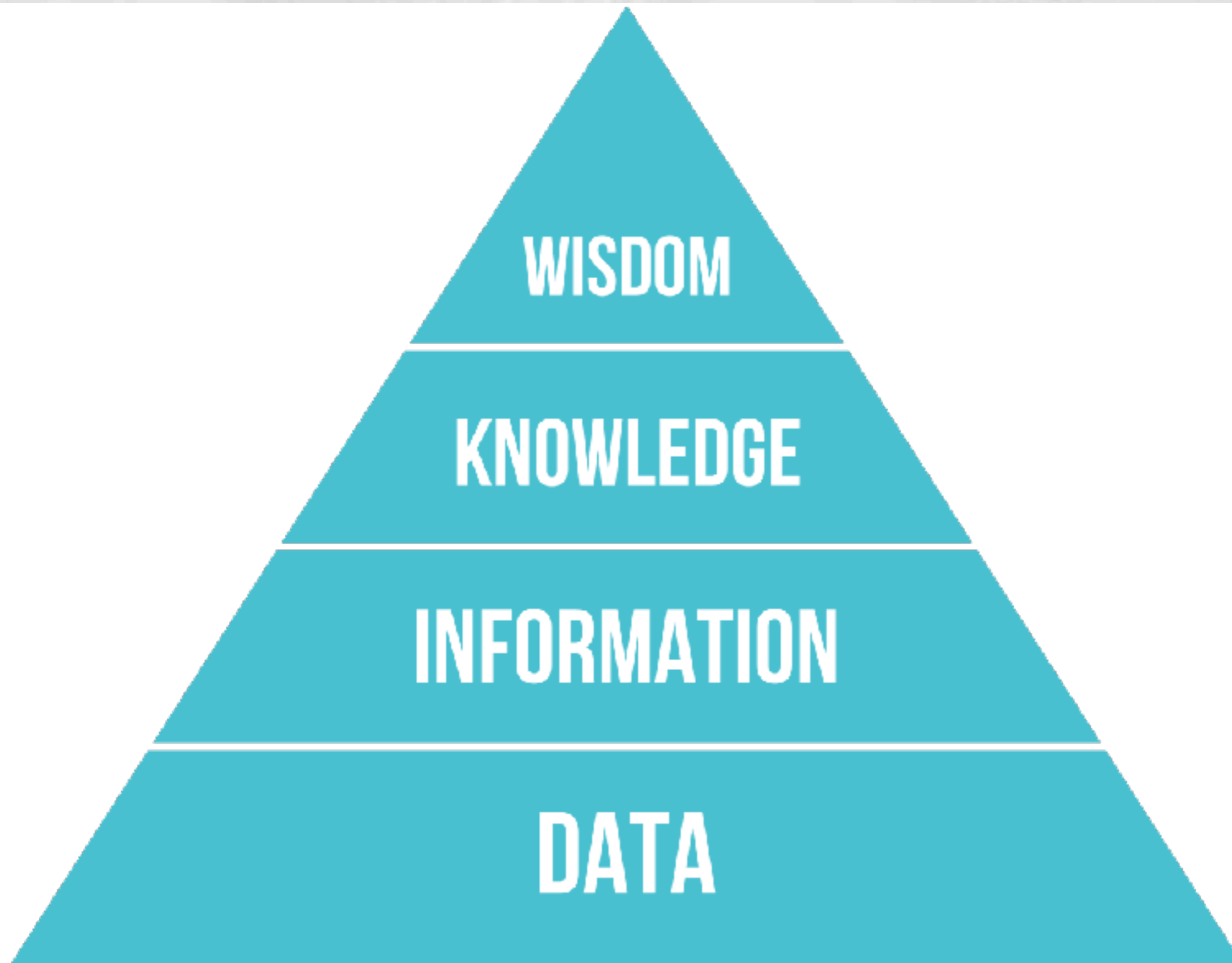
GFBio e.V.

✉️ ivo@gfbio.org

🐦 [@tigroumaniac](https://twitter.com/tigroumaniac)



DATA



https://upload.wikimedia.org/wikipedia/commons/0/06/DIKW_Pyramid.svg

90 Billion €

<http://www.dfg.de/sites/foerderatlas2018>

6.8 Billion Euro of third party funding was shared between Universities in Germany in 2015

<http://www.dfg.de/sites/foerderatlas2018>

Förderatlas 2018

KARTENANSICHT

PUBLIKATION

THEMEN

DOWNLOADS

KONTAKT

DFG Deutsche
Forschungsgemeinschaft

Förderatlas 2018

Kennzahlen zur öffentlich finanzierten
Forschung in Deutschland

ALSO E-PAPER ÖFFNEN

.. for discovering and reusing multiple data sources

80%

Mons, B. et al., doi:10.3233/ISU-1704824

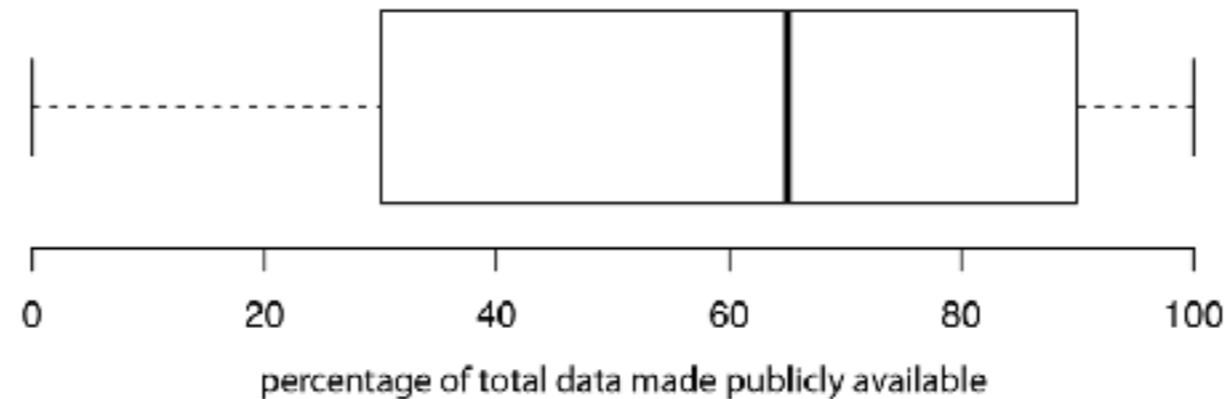
dark data

A dark, high-angle photograph of a server room. The room is filled with rows of server racks, creating a grid-like pattern of light and shadow. Two people are visible in the center, looking at a monitor on a desk. The overall atmosphere is mysterious and technical.

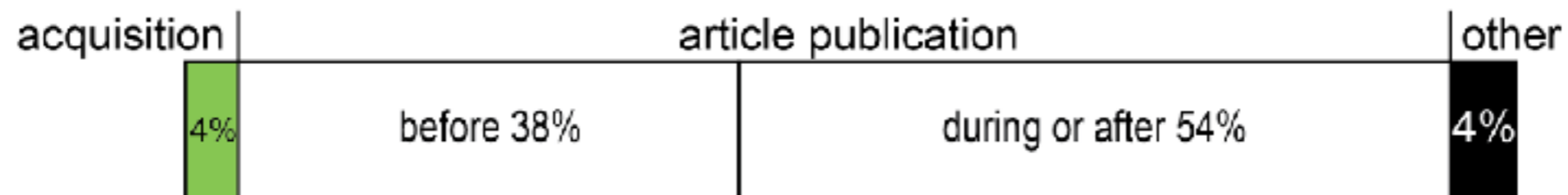
[...] data that has never been published or otherwise made available to the rest of the scientific community.

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

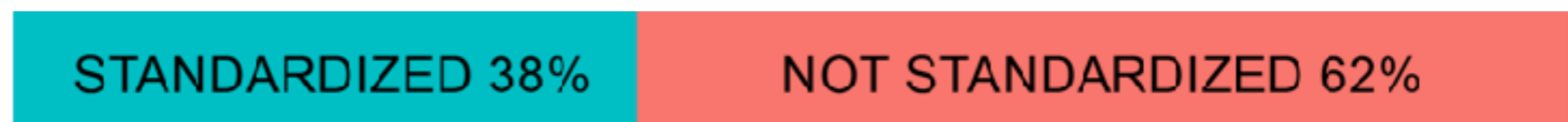
- Not all the data is archived/published



- Data archiving/publication is tied to journal articles

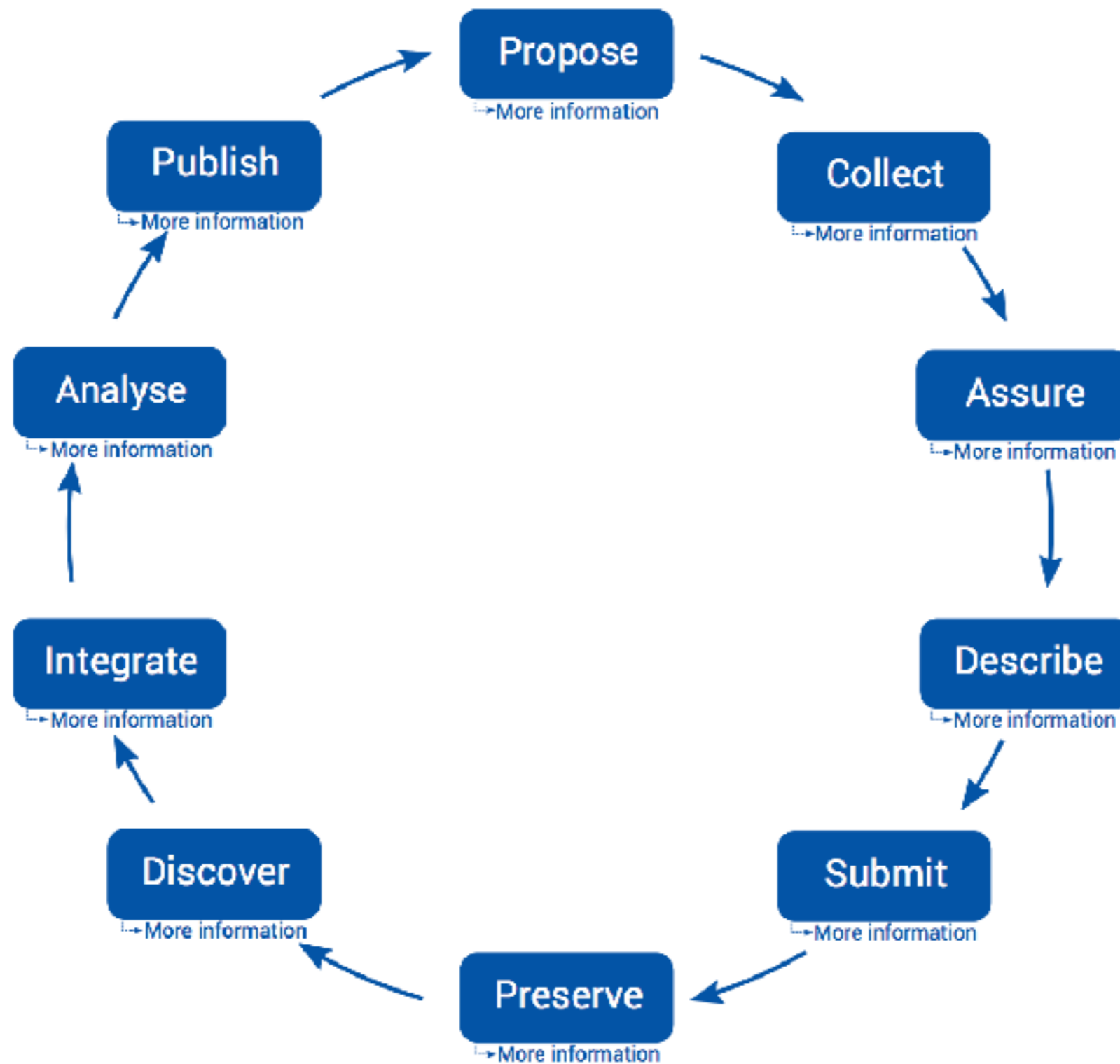


- Contextual data is rarely standardized



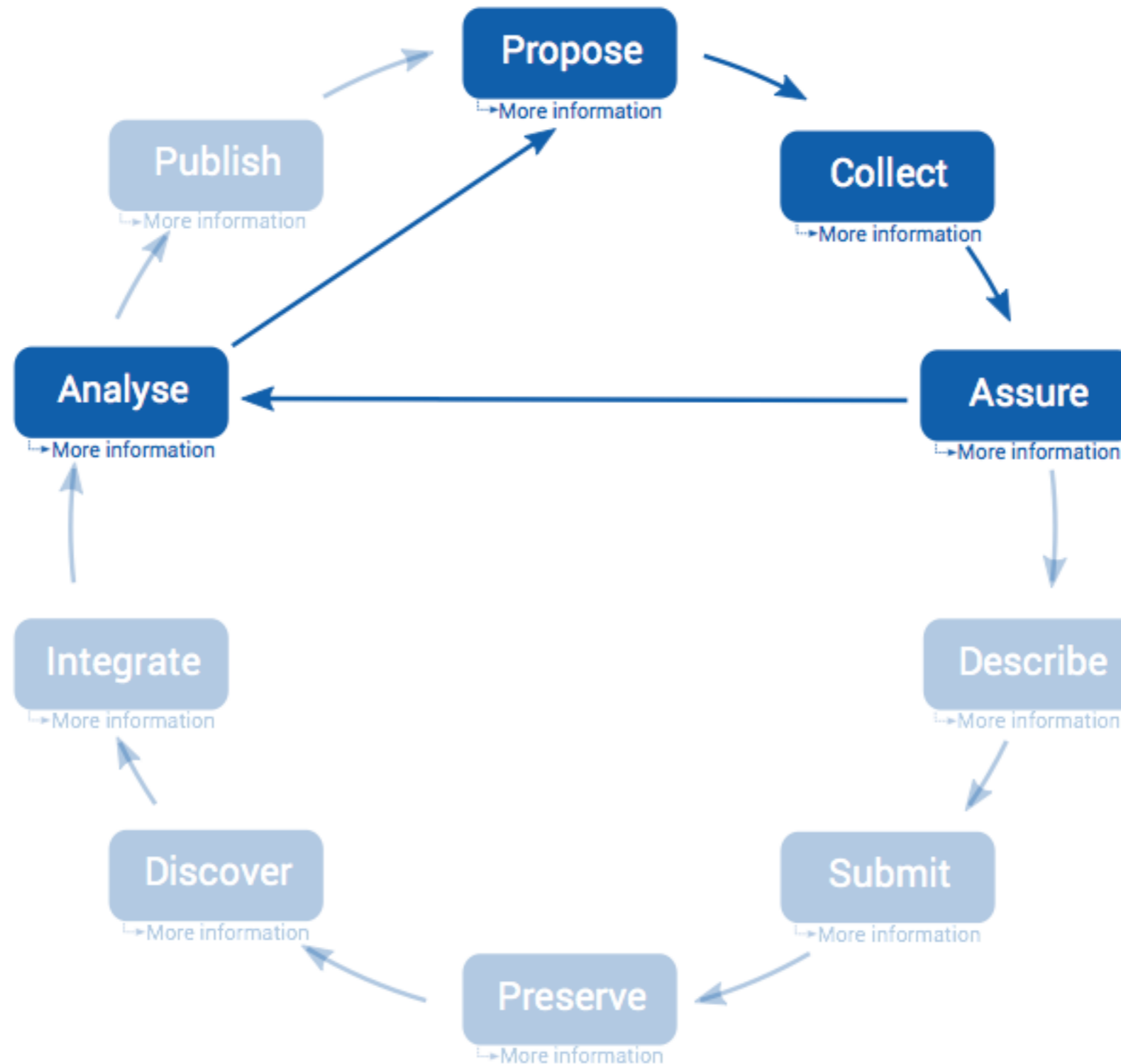
*based on user survey

Data Lifecycle



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

Data Lifecycle



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

- Data is constantly growing in size and complexity
- Datasets are hard to find and even harder to compare
- No suitable credit for (good) data publication

A lot of primary data gets effectively lost!

Community response

F
Findable



A
Accessible



I
Interoperable



R
Reusable



www.nature.com/scientificdata

SCIENTIFIC DATA

OPEN

SUBJECT CATEGORIES

- » Research data
- » Publication characteristics

Comment: The FAIR Guiding Principles for scientific data management and stewardship

Mark D. Wilkinson *et al.*[#]

There is an urgent need to improve the infrastructure supporting the reuse of scholarly data. A diverse set of stakeholders—representing academia, industry, funding agencies, and scholarly publishers—have come together to design and jointly endorse a concise and measureable set of principles that we refer to as the FAIR Data Principles. The intent is that these may act as a guideline for those wishing to enhance the reusability of their data holdings. Distinct from peer initiatives that focus on the human scholar, the FAIR Principles put specific emphasis on enhancing the ability of machines to automatically find and use the data, in addition to supporting its reuse by individuals. This Comment is the first formal publication of the FAIR Principles, and includes the rationale behind them, and some exemplar implementations in the community.

Received: 10 December 2015

Accepted: 12 February 2016

Published: 15 March 2016

Findable
Accessible
Interoperable
Reusable

Wilkinson, et al., Scientific Data, 2016
<http://doi.org/10.1038/sdata.2016.18>

Box 2 | The FAIR Guiding Principles

To be Findable:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

To be Accessible:

- A1. (meta)data are retrievable by their identifier using a standardized communications protocol
 - A1.1 the protocol is open, free, and universally implementable
 - A1.2 the protocol allows for an authentication and authorization procedure, where necessary
- A2. metadata are accessible, even when the data are no longer available

To be Interoperable:

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (meta)data use vocabularies that follow FAIR principles
- I3. (meta)data include qualified references to other (meta)data

To be Reusable:

- R1. meta(data) are richly described with a plurality of accurate and relevant attributes
 - R1.1. (meta)data are released with a clear and accessible data usage license
 - R1.2. (meta)data are associated with detailed provenance
 - R1.3. (meta)data meet domain-relevant community standards

Wilkinson, et al., *Scientific Data*, 2016 <http://doi.org/10.1038/sdata.2016.18>

- a standard
- equal to open data
- a quality, but a quantity
- only for humans or only for machines
- only for life sciences
- equal to RDF, Linked Data, or Semantic Web

B. Mons et al., doi:10.3233/ISU-170824

- Good scientific practice
- Career boost
 - article acceptance
 - data reuse & citation
 - proposal funding
 - compatibility with future infrastructures
- Career opportunities as a data scientist, manager, steward, custodian, librarian, etc.
- Keep your research legal (i.e. avoid biopiracy)



Image: <https://www.flickr.com/photos/ciat/3887465932>

Biopiracy happens when researchers or research organisations take biological resources without official sanction, largely from less affluent countries or marginalised people.

<http://theconversation.com/biopiracy-when-indigenous-knowledge-is-patented-for-profit-55589>

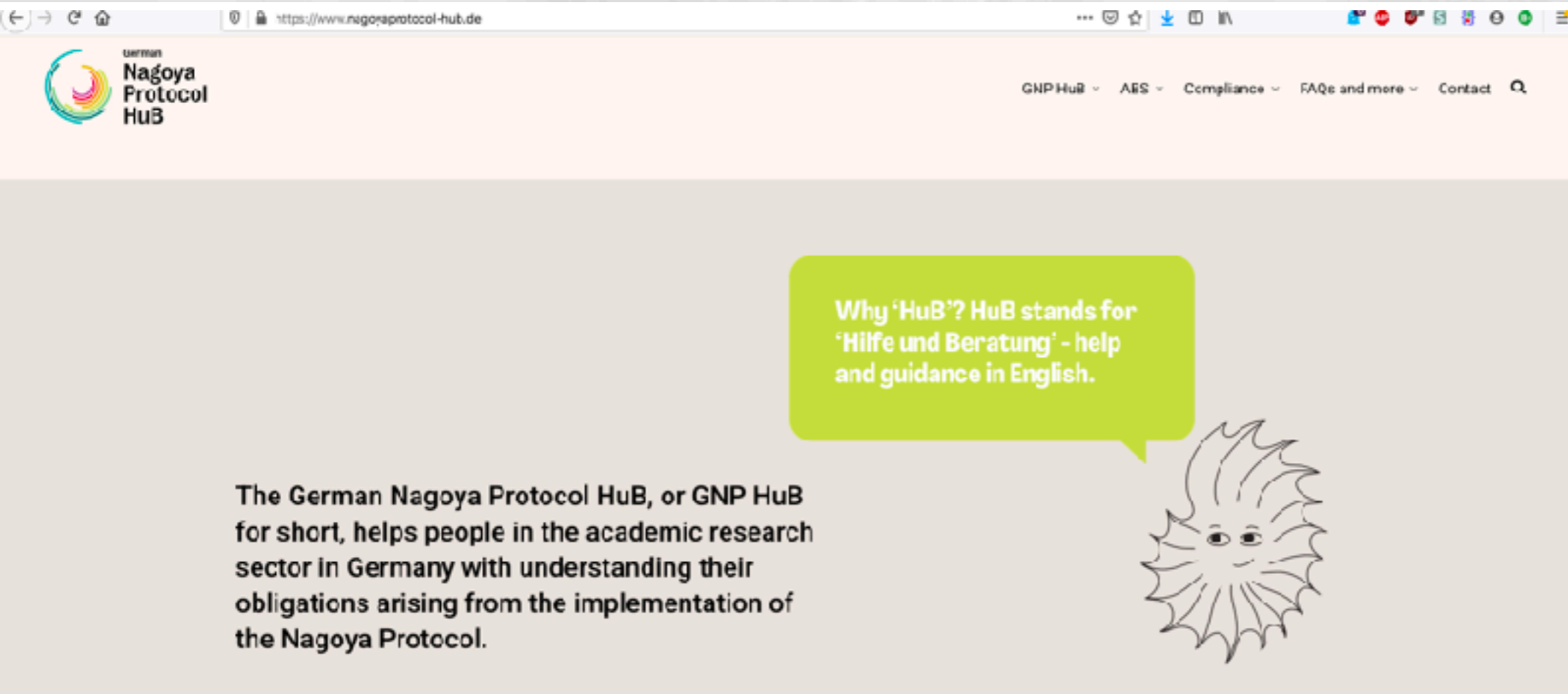


The Nagoya Protocol on Access and Benefit-sharing

A transparent legal framework for the fair and equitable sharing of benefits arising out of the utilization of genetic resources.

<https://www.cbd.int/abs/about/>

Nagoya Hub



The screenshot shows the homepage of the German Nagoya Protocol HuB. The browser address bar displays 'https://www.nagoyaprotocol-hub.de'. The website header includes the logo 'German Nagoya Protocol HuB' and a navigation menu with items: 'GNP HuB', 'AES', 'Compliance', 'FAQs and more', and 'Contact'. The main content area features a green speech bubble with the text: 'Why 'HuB'? HuB stands for 'Hilfe und Beratung' - help and guidance in English.' Below this, a paragraph states: 'The German Nagoya Protocol HuB, or GNP HuB for short, helps people in the academic research sector in Germany with understanding their obligations arising from the implementation of the Nagoya Protocol.' To the right of the text is a stylized drawing of a sun with a face.

About the German Nagoya Protocol HuB

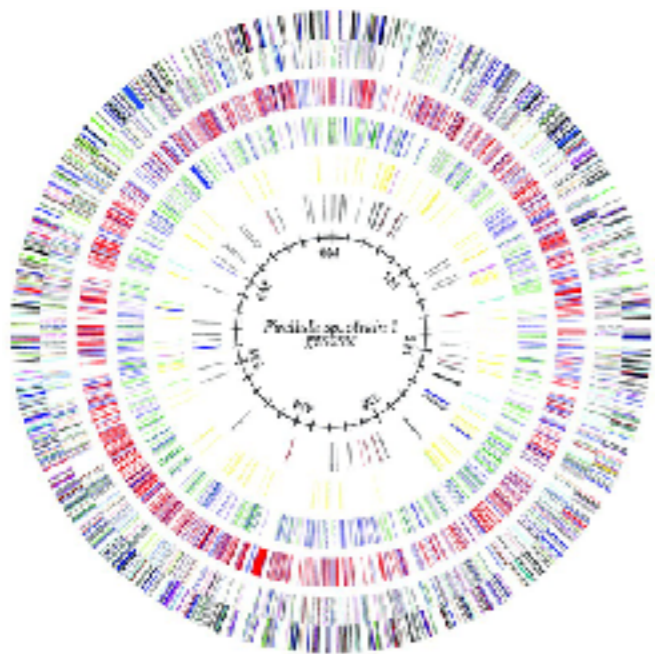
Behind the scenes – The Leibniz Institute-German Collection for Microorganisms and Cell Cultures (DSMZ) is partnering with the Consortium of German Natural Science Collections (DNFS), the German Life Sciences Association (VBIO) and the Leibniz Research Alliance for Biodiversity (VRB) on the German Nagoya Protocol HuB.

You can find more information about the project and our activities under [about us](#) and in the [library](#).



<https://www.nagoyaprotocol-hub.de/>

METADATA & STANDARDS



Glöckner et al., 2003

DATA

land use nitrate salinity host relationship cell
size motility calcium perturbation 16S sulfide bromide
exoenzymes chemotaxis biofilm products antibiotics
metabolism halophily magnesium substrate spectrum isolation
oxygen pathogenicity light phosphate carbon
classification genome organic matter
pigmentation ammonium sulfate C/N ratio
gram stain ph CO2 cultivation temperature

Courtesy: Boyke Bunk, DSMZ

METADATA

contextual data

MixS **M**inimal **I**nformation about any(**x**) **S**equence

<http://gensc.org/mixs/>

developed by:



supported by:

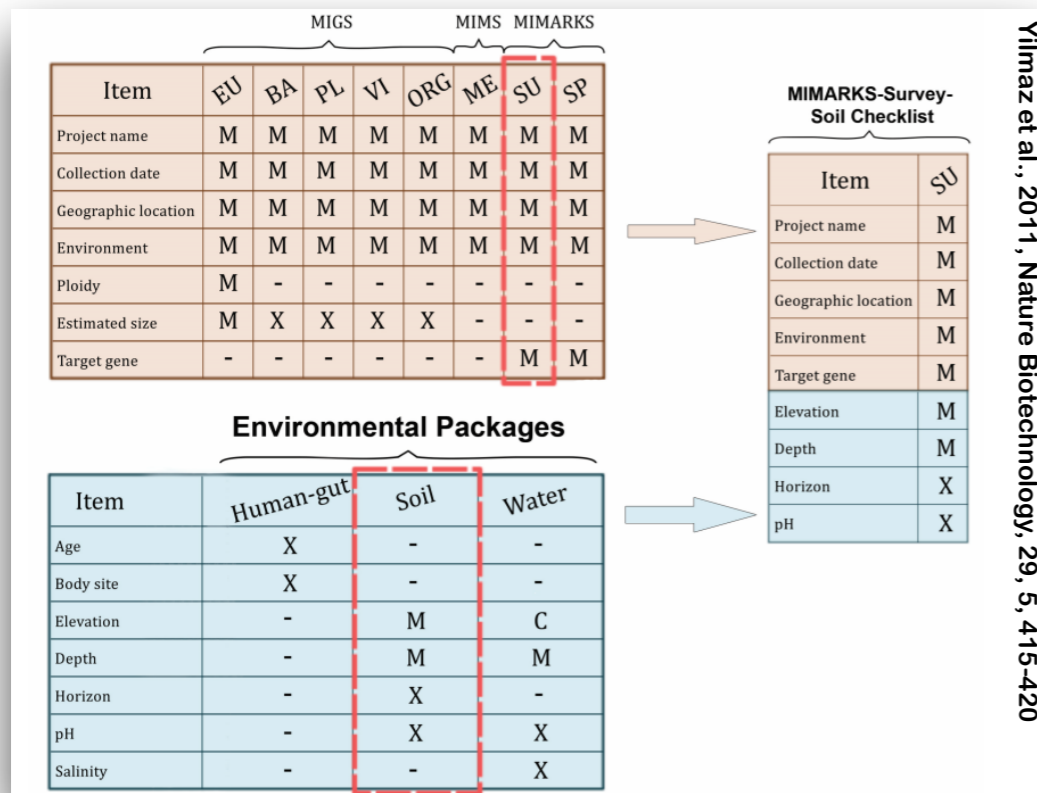
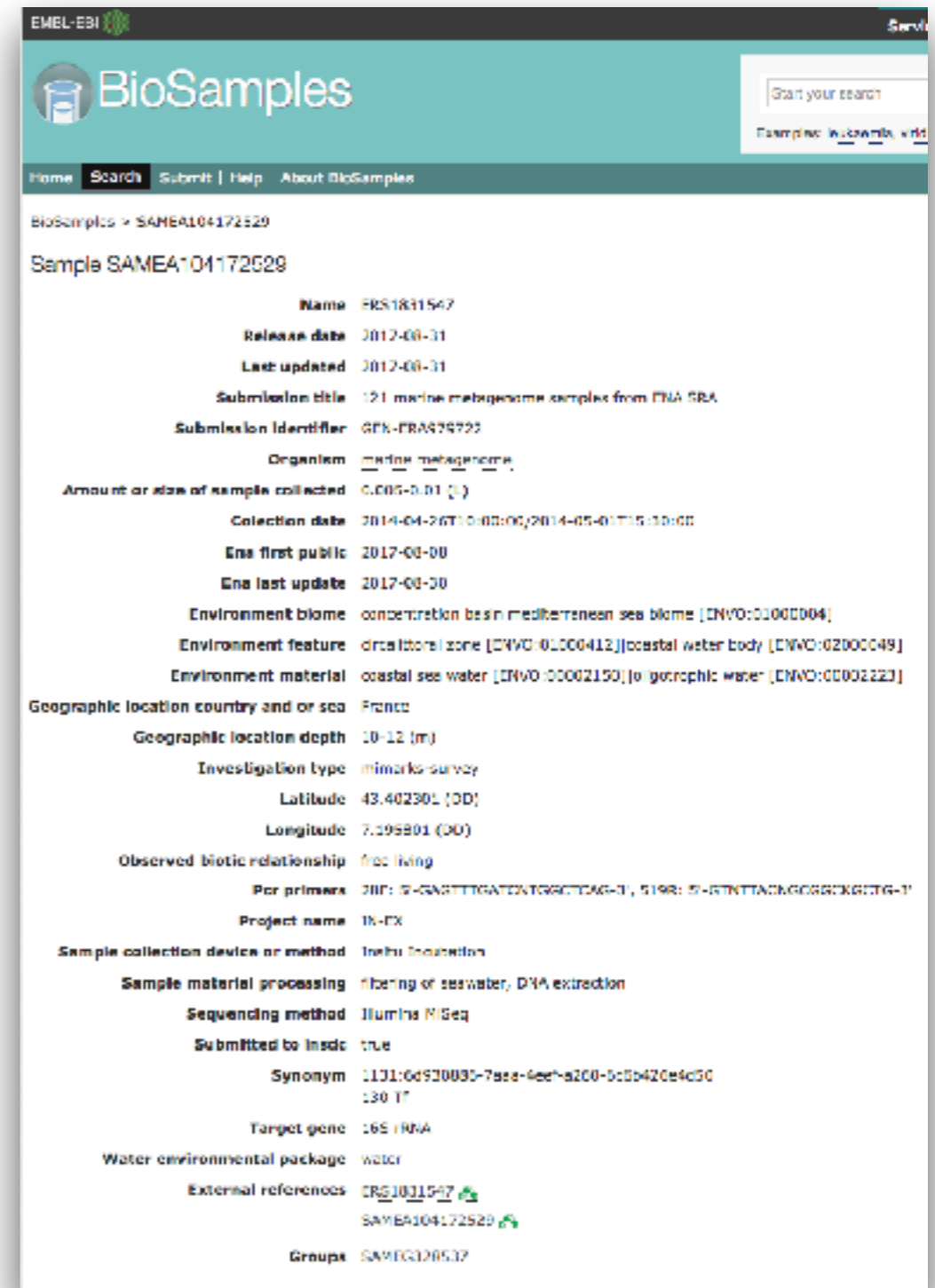


MIGS/MIMS - Field et al., Nature Biotechnology 26, 541 - 547 (2008)

MIMARKS & MixS - Yilmaz et al., Nature Biotechnology 29, 415–420 (2011)

MixS - Yilmaz et al., The ISME Journal 5, 1565–1567 (2011)

MiXS Minimal Information about any(x) Sequence

BioSamples

Home Search Submit Help About BioSamples

BioSamples > SAMEA104172529

Sample SAMEA104172529

- Name: FR51831547
- Release date: 2012-08-31
- Last updated: 2012-08-31
- Submission title: 121 marine metagenome samples from FNA SRA
- Submission identifier: GEN-FR51831547
- Organism: marine metagenome
- Amount or size of sample collected: 0.005-0.01 (g)
- Collection date: 2014-04-26T10:00:00;2014-05-01T15:10:00
- Ens first public: 2012-08-30
- Ens last update: 2012-08-30
- Environment biome: concentration basin mediterranean sea biome [ENVO:01000004]
- Environment feature: circaliboral zone [ENVO:01000412];coastal water body [ENVO:02000049]
- Environment material: coastal sea water [ENVO:00002150];oligotrophic water [ENVO:00002223]
- Geographic location country and/or sea: France
- Geographic location depth: 10-12 (m)
- Investigation type: mimarks-survey
- Latitude: 43.402301 (DD)
- Longitude: 7.195801 (DD)
- Observed biotic relationship: free living
- PCR primers: 318F: G-CAGTTTGATGATGAGTCAAG-3', 519R: G-ATTTTACAGCGGCGKGGTGG-1'
- Project name: IN-EX
- Sample collection device or method: In situ incubation
- Sample material processing: filtering of seawater, DNA extraction
- Sequencing method: Illumina MiSeq
- Submitted to insdc: true
- Synonym: 1131:6d520080-7acc-4eeb-a200-5c5b420e4c50;130 TT
- Target gene: 16S rRNA
- Water environmental package: water
- External references: FR51831547, SAMEA104172529
- Groups: SAMEA104172529

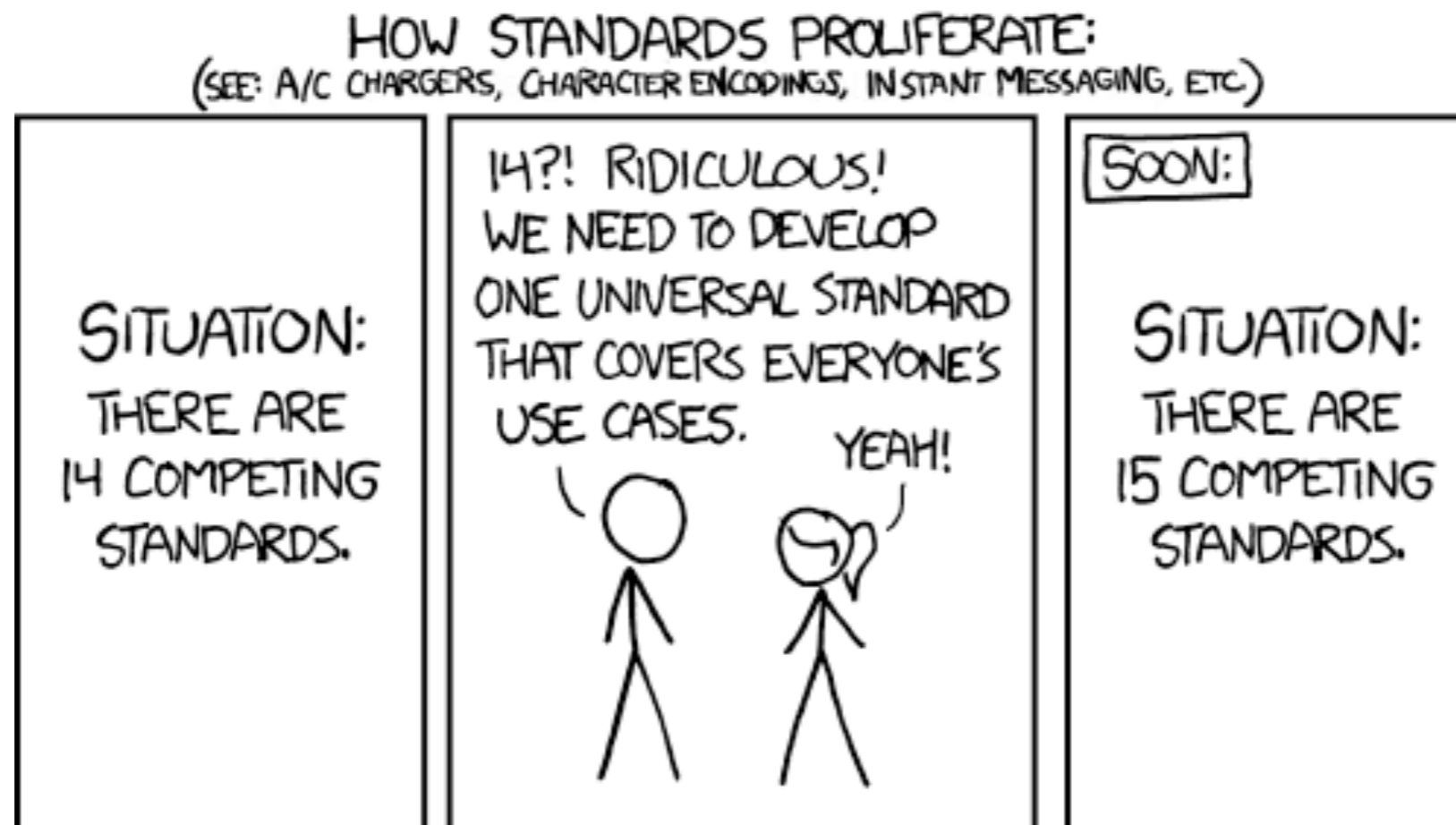
<http://gensc.org/mixs/>

MIGS/MIMS - Field et al., Nature Biotechnology 26, 541 - 547 (2008)
 MIMARKS & MiXS - Yilmaz et al., Nature Biotechnology 29, 415-420 (2011)
 MiXS - Yilmaz et al., The ISME Journal 5, 1565-1567 (2011)

Standards for biodiversity & ecology:

- Darwin Core (Wieczorek et al. 2012)
- ABCD (Holetschek et al. 2012)
- GGBN (Dröge et al. 2016)
- EML (<https://knb.ecoinformatics.org/#tools/eml>)

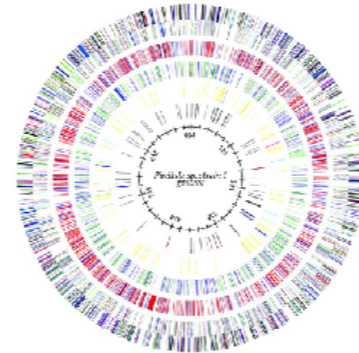
- Insufficient (acquisition) tooling
- Insufficient training & support



<http://xkcd.com/927/>

ARCHIVAL & PUBLICATION

Findable
Accessible
Interoperable
Reproducible



land use nitrate salinity host relationship cell
size motility calcium perturbation 16S sulfide bromide
exoenzymes chemotaxis biofilm products antibiotics
metabolism halophily magnesium substrate spectrum isolation
oxygen pathogenicity light phosphate carbon
classification genome organic matter
pigmentation ammonium sulfate C/N ratio
gram stain ph CO2 cultivation temperature



dedicated, long-term archives




Examples: BN000065, histone


[Search](#)
[Advanced Sequence](#)

[ENA](#) > [Submit and update](#)

Submitting and updating data

We offer a number of services through which data (including updates) can be submitted to the European Nucleotide Archive (ENA). These technologies provide options appropriate for the scale and frequency of submission, the expertise and capacity of the submitter and the nature of the data to be transferred. The choices below lead users most directly to the appropriate submission route.

 [Submit to ENA](#)

 [Email ENA Helpdesk](#)

Programmatic submissions

Most types of submissions can also be made programmatically ... [more information.](#)

Submit & Update

- ▶ [Data formats](#)
- [Taxon ID requests](#)
- [Uploading data files](#)
- ▶ [Reads](#)
- ▶ [Sequences](#)
- ▶ [Genome assembly submissions](#)
- [Taxonomy](#)
- ▶ [Sample checklists](#)
- [Environmental](#)
- [Epigenomic](#)
- [Species BARCODE](#)
- [Metadata model](#)
- [Register submission account](#)
- ▶ [Programmatic XML submissions](#)

Please specify the release date of your study:
This is when your study will be made public.

Please provide a short name for the study

Please provide a short descriptive title

Please provide an abstract to describe your study

For genome assembly projects only: Is this a genome assembly project?
PLEASE ANSWER WITH YES IF YOU HAVE A GENOME ASSEMBLY.

Yes
 No

Please provide PubMed IDs of publications related to your study
(numeric value)

PubMed IDs

Please provide attributes to add a deep description of your study

Tag	Value
<input type="text"/>	<input type="text"/>

Please select the checklist attributes you would like to include with each sample. Recommended attributes can be unselected from within the corresponding attribute group on the left-hand side panel. You may also add custom attributes.

 internal environment local environment conditions non-sample term sample collection organism characteristics concentration measurement host details other host description geography User Attributes

12 of 78 attributes selected

Please complete any fields that you would like to apply to all samples. This will act as a template for the rest of the samples.

Template Basic Details

Unique Name Prefix:

* Title:

Description:

Organism Details

If your organism is not found please go [here](#) and email datasubs@ebi.ac.uk with the required details listed on the page in order for us to request a taxon id for your organism.

Search:

* Tax Id:

* Scientific Name:

Common Name:

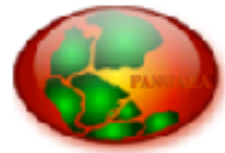
non-sample term

* investigation type:

* project name:

* sequencing method:

sample collection



PANGAEA®
Data Publisher for Earth & Environmental Science

Logged in as **ikostadi** ([log out](#), [profile](#))

Submit Data to PANGAEA

Welcome to the PANGAEA data submission system. Any data from earth and life sciences are accepted. We highly appreciate you archiving and publishing your data with PANGAEA.

- ✔ **Benefits.** Published data are fully citable and can be cross-referenced with journal articles – [read more...](#)
- ✔ **Data preparation and quality control.** We will be in direct contact with you during preparation and archiving of your data – [read more...](#)
- ✔ **Costs.** PANGAEA can be used free of charge. Nevertheless, we appreciate any possible financial support – [read more...](#)

When you start the data submission process below, you will be redirected to the PANGAEA issue tracker that will assist you in providing metadata and uploading data files. Any communication with our editors will go through this issue tracker.

SUBMIT YOUR DATA

[Further details about data submission to PANGAEA](#) – if you have any comments on the data submission process, please [contact us](#).

[Contact](#)

Easy for the user,
not so much for the curator.

Project **PANGAEA Data Archiving & Publication**

Issue Type **Data Submission**

Summary*
The summary (subject) is used as identifier in the further communication.

Author(s)*
Please, enter the author(s) (the principal investigators) for the data set(s) you want to submit.
One author per line, example: Smith, Joe Peter

Title
The title should ideally reflect what has been measured, observed, or calculated, when, where, and how.

Description
ABSTRACT and/or further details describing the data.

Keywords
Separate keywords by comma or semicolon.

Attachment

Drop files here to attach them

or

For larger files leave a corresponding note in the description --DATA FILE(S) ARE REQUIRED
 For data submissions, read our format guide.

License*
General information on used licenses can be found on the [Creative Commons](#) license pages. If you need help to choose the correct license for your dataset, you can use the [following page](#).

Labels
Begin typing to find and create labels or press down to select a suggested label.
Content of the data submission, e.g. PROJECT, Institute, etc.
Labels have to be one word!

Data used/published in the following article/manuscript
Please, specify any references to articles or submitted manuscripts related to this data submission.
Copy/paste the full citation, preferably with a [Digital Object Identifier \(DOI name\)](#)

- Different archives have different submission and curatorial workflows, requirements, response times.
- Your time is limited. Do you prefer to do analysis or submissions?
- Incentives (i.e. credits) for high-quality, FAIR data often unclear.



www.gfbio.org



A sustainable, service-oriented network of infrastructures and experts.

Funded by:





FAIR • Research • Data

Biodiversity, Ecology & Environmental Science

Enter a search term...

FIND DATA



Plan

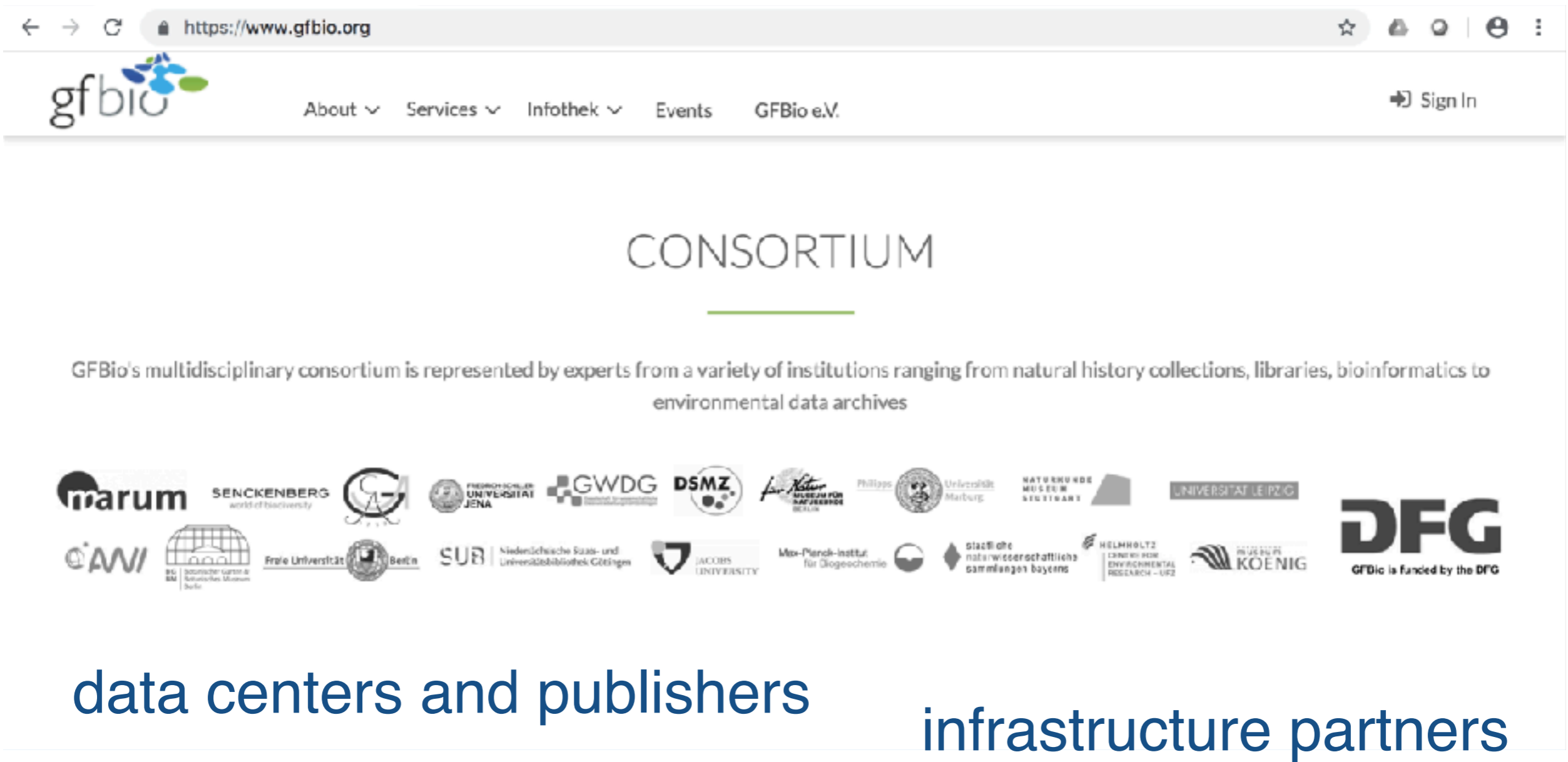


Submit



Visualize

Network of experts



The screenshot shows the GFBio website with the following content:

- Browser address bar: <https://www.gfbio.org>
- Navigation menu: About, Services, Infothek, Events, GFBio e.V., and a Sign In button.
- Section title: CONSORTIUM
- Text: "GFBio's multidisciplinary consortium is represented by experts from a variety of institutions ranging from natural history collections, libraries, bioinformatics to environmental data archives"
- Logos of partner institutions: marum, SENCKENBERG world of biodiversity, GARD, FRIEDRICH-SCHILLER UNIVERSITÄT JENA, GWDG, DSMZ, Leibniz-Museum für Naturkunde Berlin, Philipps-Universität Marburg, NATURHISTORISCHES MUSEUM SENFTENAU, UNIVERSITÄT LEIPZIG, ANI, Staatliches Naturhistorisches Museum Berlin, Freie Universität Berlin, SUB | Niedersächsische Staats- und Universitätsbibliothek Göttingen, JACOBS UNIVERSITY, Max-Planck-Institut für Geochemie, Staatliche Naturwissenschaftliche Sammlungen Bayerns, HELMHOLTZ-Zentrum für Umweltforschung - UFZ, and DFG (German Research Foundation).

data centers and publishers

infrastructure partners

user-representatives



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.



TERMINOLOGY SERVICE

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

- A good DMP should cover:
 - Data acquisition
 - Quality assurance
 - Intermediate handling and storage
 - Long-term archiving
 - Analysis
 - Publication (open-access, licensing)

- **Basic information:** project title, contact information, motivation for data collection
- **Information on data:** type, format, volume, collection standards, methodologies, quality assurance
- **Documentation and metadata:** readability and interpretability of data, metadata standards
- **Ethical and legal compliance:** agreement on preservation/sharing conditions, sensitive data, intellectual property
- **Storage and backup plan:** responsibility, data recovery, access for collaborators, security

<https://www.gfbio.org/training/materials/data-lifecycle/plan>

- **Preservation:** selection of data, foreseeable future use, time and location for preservation, costs
- **Data sharing and publication:** modalities, conditions, persistent identifiers
- **Responsibilities:** implementation, roles and responsibilities for each activity, ownership agreement
- **Resources:** need for additional hardware/software or expertise for training, efforts and costs for data management and data archiving

<https://www.gfbio.org/training/materials/data-lifecycle/plan>



About ▾

Data ▾

Training ▾

Support ▾

News ▾

Contact

GFBio e.V.

Sign In

Welcome to the

GFBio Data Management Plan Tool!

- ✓ Collect information about your project
- ✓ Complete your DMP checklist
- ✓ Get GFBio DMP support




Get started


The GFBio Data Management Planning Tool supports you in preparing your custom DMP. It helps you think about the most important questions concerning data management as early as possible. Collect information about your project, fill in the DMP checklist and send us your DMP support request. We will support you in optimizing your data management and finalizing your data management plan.

Learn more: [How to create a data management plan \(DMP\)?](#)

DMP Support

← → ↻ 🏠 <https://www.gfbio.org/data/plan/dmpt> 90% 🔔 ⭐ 📄 🗑️ 👤

 About ▾ Data ▾ Training ▾ Support ▾ News ▾ Contact GFBio e.V. Welcome Ivaylo! [Sign Out](#)

 **1. General Project Information** 2. Data Collection 3. Documentation and Metadata 4. Ethics and Legal Compliance 5. Preservation and Sharing

What is the official name of your research project? *

TEST

Please select a category:

Other ▾

Is your research data reproducible? ⓘ

One-time observation Repeatable experiments Time series

Add additional information (e.g. data reproduction might cause high costs or a lot of effort).

Please specify your project type. ⓘ

<input checked="" type="radio"/> Field Work	<input type="radio"/> Simulation
<input type="radio"/> Observational	<input type="radio"/> Assimilation
<input type="radio"/> Experimental	<input type="radio"/> Modelling
<input checked="" type="radio"/> Laboratory	<input type="radio"/> Other

Provide your project abstract or describe your work and the data involved.

Data Management Plan Tool

Send a DMP support request to GFBio, download your DMP or save it to your private dashboard.



Request Data Management Plan Support

Send Request



Download PDF-File

Download



Save Data Management Plan

Save



Finish Wizard

Finish

Data Submission

SIGN UP

SIGN IN



Data Submissions

Long-term data archival & publication services for Biodiversity,
Ecology & Environmental Science

SUBMIT YOUR DATA !

Already using our service? [Sign in](#)

<https://submissions.gfbio.org>

Data Submission

[My Submissions](#) [Create Submission](#)

Title

Enter a title for your dataset

Description

Describe your dataset.

Upload Data (optional)

Try **dropping** some files here, or **click** to select files to upload.

Data URL (optional)

Link to your data, e.g. cloud storage

Contributors (optional)

Contributors List

[+ add contributor](#)

Target Datacenter (optional)

GFBio Data Centers - our curators will suggest the appropriate one(s)

Info

[Do you need Help ?](#)

License

[CC BY 4.0](#) [change](#)

Legal Requirements

- Nagoya Protocol
- IUCN Red List of Threatened Species
- Sensitive Personal Information
- Uncertain

Metadata Templates [?](#)

Molecular Data Template:

- [CSV Template](#)
- [Template Description](#)

Biodiversity, Ecological and Collection Data Template:

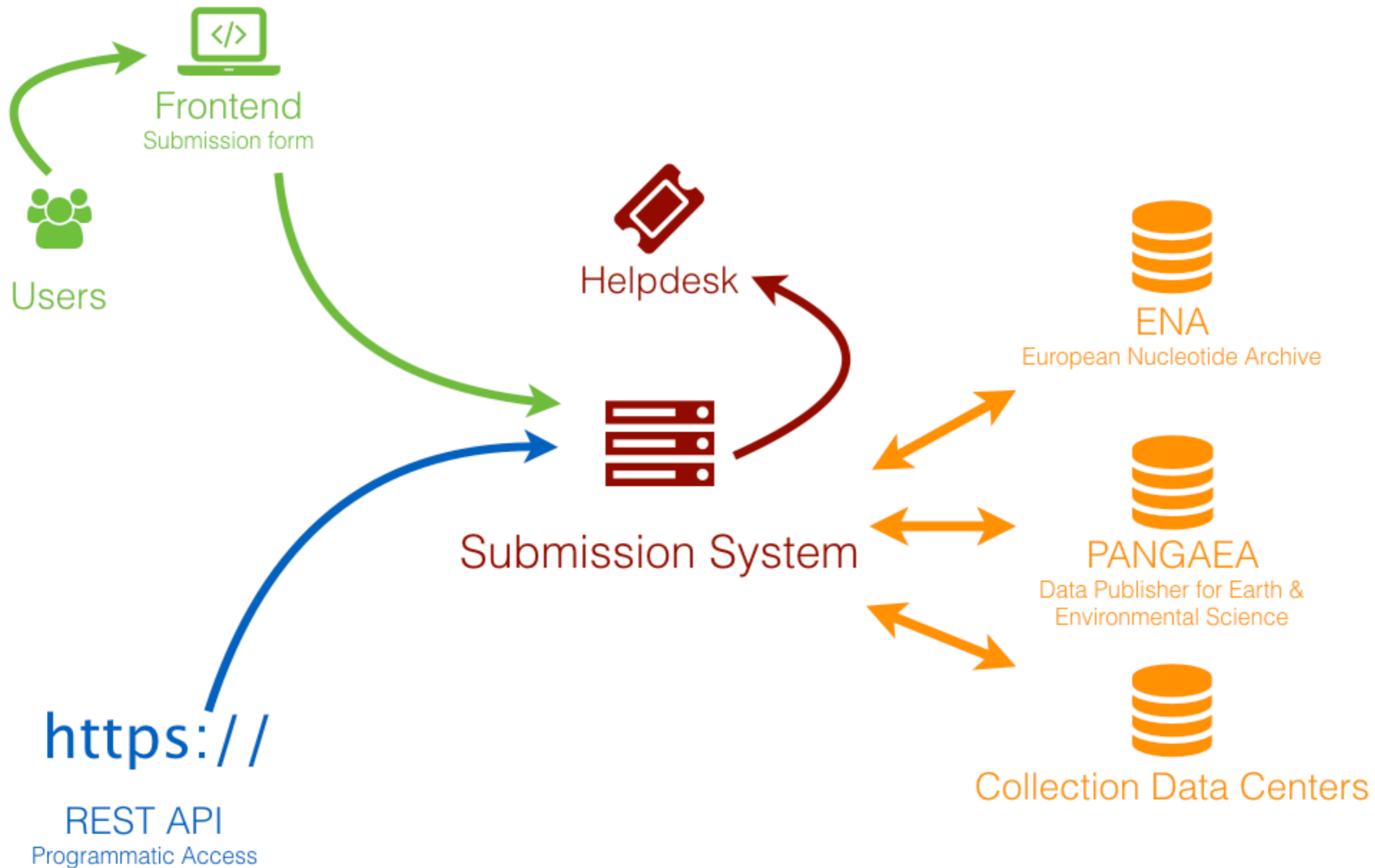
- [CSV Template](#)
- [Template Description](#)

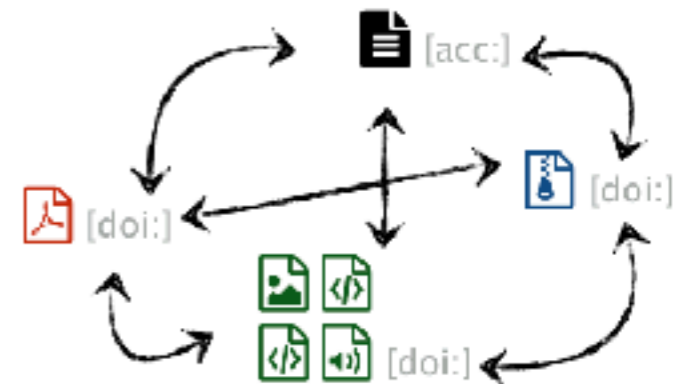
Embargo Date

13 October 2021

[Change embargo date](#)

Data Submission

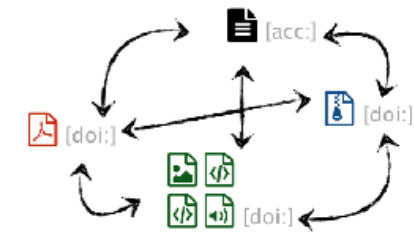




Added value at a glance:

- Single-point of contact - data is distributed to data centers and interlinked
- Expert support for metadata standardization - ABCD, DwC, **MixS**, **ENVO**
- Manual and programmatic (API) operation
- Integration with local RDM systems

Data publication & citation



europaepmc.org/abstract/PMC/PMC6265507

Europe PMC

About Tools Developers Help **β Explore the beta version** Europe PMC plus

Search worldwide, life-sciences literature

E.g. "breast cancer" HER2 Smith J

Search Advanced Search

Environmental Drivers of Free-Living vs. Particle-Attached Bacterial Community Composition in the Mauritania Upwelling System. (PMCID:PMC6265507)

Abstract Citations Related Articles Data BioEntities External Links

Bachmann J¹, Heimbach T¹, Hassenrück C¹, Kopprio GA¹, Iversen MH², Grossart HR³, Gärdes A¹

Affiliations >

Frontiers in Microbiology (23 Nov 2018, 9:2836)

Type: research-article, Journal Article
DOI: 10.3389/fmicb.2018.02836

Abstract

Saharan dust input and seasonal upwelling along North-West Africa provide a model system for studying microbial processes related to the export and recycling of nutrients. This study offers the first molecular characterization of prokaryotic particle-attached (PA; >3.0 μm) and free-living (FL; 0.2-3.0 μm) players in this important ecosystem during August 2016. Environmental drivers for alpha-diversity, bacterial community composition, and differences between FL and PA fractions were identified. The ultra-oligotrophic waters off Senegal were dominated by Cyanobacteria while higher relative abundances of Alphaproteobacteria, Bacteroidetes, Verrucomicrobia, and Planctomycetes (known particle-degraders) occurred in the upwelling area. Temperature, proxy for different water masses, was the best predictor for changes in FL communities. PA community variation was best explained by temperature and ammonium. Bray Curtis dissimilarities between FL and PA were generally very high and correlated with temperature and salinity in surface waters. Greatest similarities between FL and PA occurred at the deep chlorophyll maximum, where bacterial substrate

Recent Activity Export

Formats

Abstract Full Text PDF

Show annotations in this abstract

- Chemicals
- Organisms

Data Citations

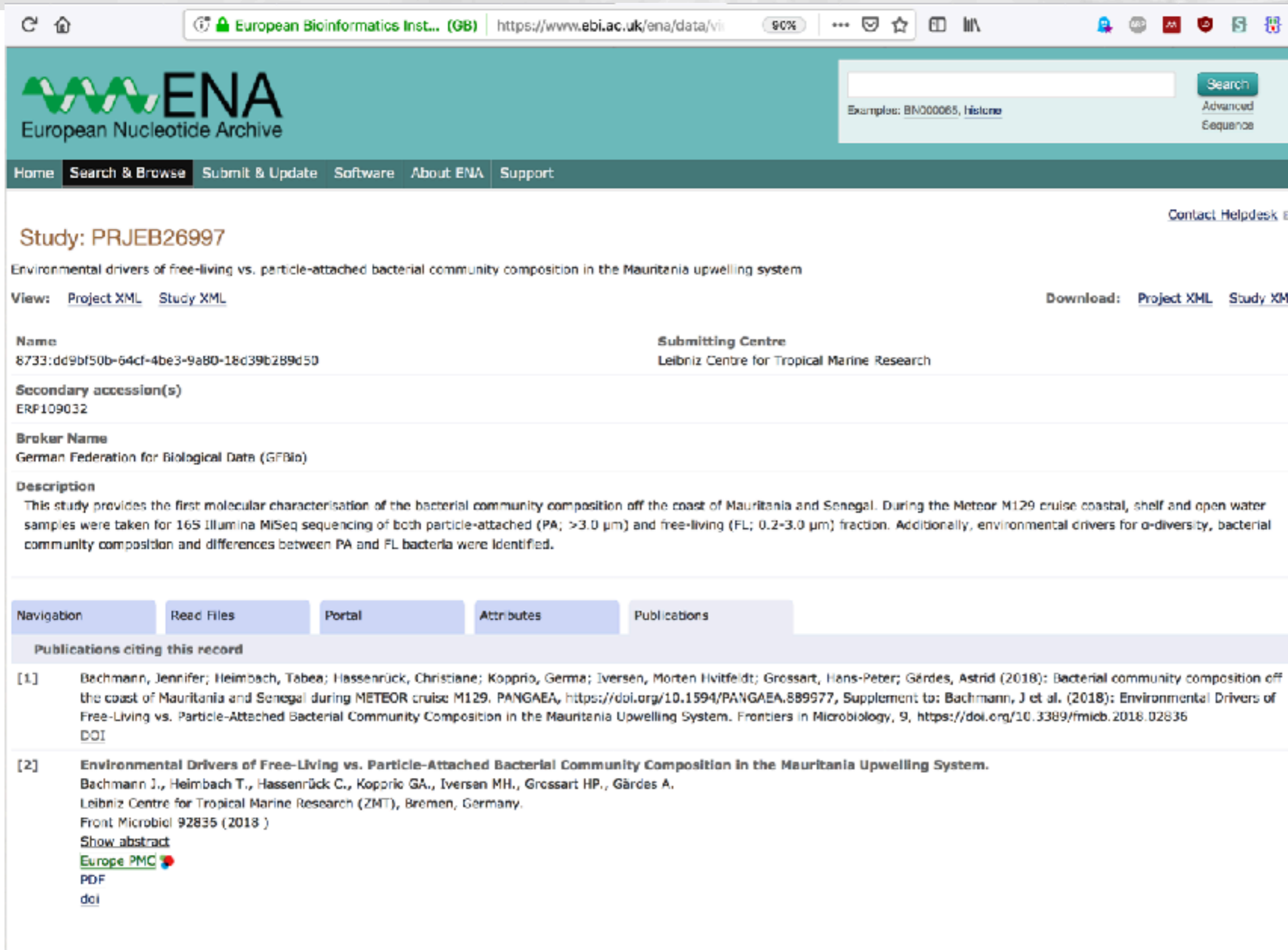
Identified 2 unique Data Citations in the full text

doi 10.1594/PANGAEA.889977 (2)

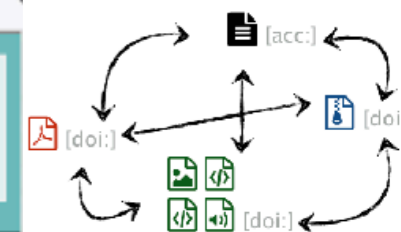
bioproject PRJEB26997 (1)

Show all items

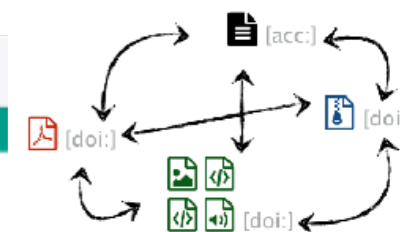
Data publication & citation



The screenshot shows the ENA website interface. At the top, there is a search bar with the text "Example: BN000065, histone" and a "Search" button. Below the search bar is a navigation menu with options: Home, Search & Browse, Submit & Update, Software, About ENA, and Support. The main content area displays the study record for PRJEB26997, titled "Environmental drivers of free-living vs. particle-attached bacterial community composition in the Mauritania upwelling system". The record includes the name "8733:dd9bf50b-64cf-4be3-9a80-18d39b289d50", the submitting centre "Leibniz Centre for Tropical Marine Research", and the secondary accession "ERP109032". The broker name is "German Federation for Biological Data (GFBio)". The description states: "This study provides the first molecular characterisation of the bacterial community composition off the coast of Mauritania and Senegal. During the Meteor M129 cruise coastal, shelf and open water samples were taken for 16S Illumina MiSeq sequencing of both particle-attached (PA; >3.0 µm) and free-living (FL; 0.2-3.0 µm) fraction. Additionally, environmental drivers for α-diversity, bacterial community composition and differences between PA and FL bacteria were identified." Below the description is a navigation bar with tabs: Navigation, Read Files, Portal, Attributes, and Publications. The "Publications" tab is active, showing a list of publications citing this record. The first publication is by Bachmann, Jennifer; Heimbach, Tabea; Hassenrück, Christiane; Koppric, Germa; Iversen, Morten Hvitfeldt; Grossart, Hans-Peter; Gärdes, Astrid (2018): Bacterial community composition off the coast of Mauritania and Senegal during METEOR cruise M129. PANGAEA, https://doi.org/10.1594/PANGAEA.889977, Supplement to: Bachmann, J et al. (2018): Environmental Drivers of Free-Living vs. Particle-Attached Bacterial Community Composition in the Mauritania Upwelling System. Frontiers in Microbiology, 9, https://doi.org/10.3389/fmicb.2018.02836 DOI. The second publication is "Environmental Drivers of Free-Living vs. Particle-Attached Bacterial Community Composition in the Mauritania Upwelling System." by Bachmann J., Heimbach T., Hassenrück C., Koppric GA., Iversen MH., Grossart HP., Gärdes A. Leibniz Centre for Tropical Marine Research (ZMT), Bremen, Germany. Front Microbiol 92835 (2018). It includes links for "Show abstract", "Europe PMC", "PDF", and "doi".



Data publication & citation



Browser address bar: <https://doi.pangaea.de/10.1594/PANGAEA.889977>

Not logged in

PANGAEA. Data Publisher for Earth & Environmental Science

SEARCH SUBMIT ABOUT CONTACT

Citation:

Bachmann, Jennifer; Heimbach, Tabea; Hassenrück, Christiane; Kopprlo, Germa; Iversen, Morten Hvitfeldt; Grossart, Hans-Peter; Gärdes, Astrid (2018): Bacterial community composition off the coast of Mauritania and Senegal during METEOR cruise M129. PANGAEA, <https://doi.org/10.1594/PANGAEA.889977>,

Supplement to: Bachmann, J et al. (2018): Environmental Drivers of Free-Living vs. Particle-Attached Bacterial Community Composition in the Mauritania Upwelling System. *Frontiers in Microbiology*, 9, <https://doi.org/10.3389/fmicb.2018.02836>

Always quote above citation when using data! You can download the citation in several formats below.

- RIS Citation
- Bibtex Citation
- Copy Citation
- Facebook
- Twitter
- Show Map
- Google Earth



Abstract:

Saharan dust input and seasonal upwelling along North-West Africa provide a model system for studying microbial processes related to the export and recycling of nutrients. This study offers the first molecular characterization of prokaryotic particle-attached (PA; >3.0 µm) and free-living (FL; 0.2–3.0 µm) players in this important ecosystem during August 2016. Environmental drivers for alpha-diversity, bacterial community composition, and differences between FL and PA fractions were identified. The ultra-oligotrophic waters off Senegal were dominated by Cyanobacteria while higher relative abundances of Alphaproteobacteria, Bacteroidetes, Verrucomicrobia, and Planctomycetes (known particle-degraders) occurred in the upwelling area. Temperature, proxy for different water masses, was the best predictor for changes in FL communities. PA community variation was best explained by temperature and ammonium. Bray Curtis dissimilarities between FL and PA were generally very high and correlated with temperature and salinity in surface waters. Greatest similarities between FL and PA occurred at the deep chlorophyll maximum, where bacterial substrate availability was likely highest. This indicates that environmental drivers do not only influence changes among FL and PA communities but also differences between them. This could provide an explanation for contradicting results obtained by different studies regarding the dissimilarity/similarity between FL and PA communities and their biogeochemical functions.

Related to:

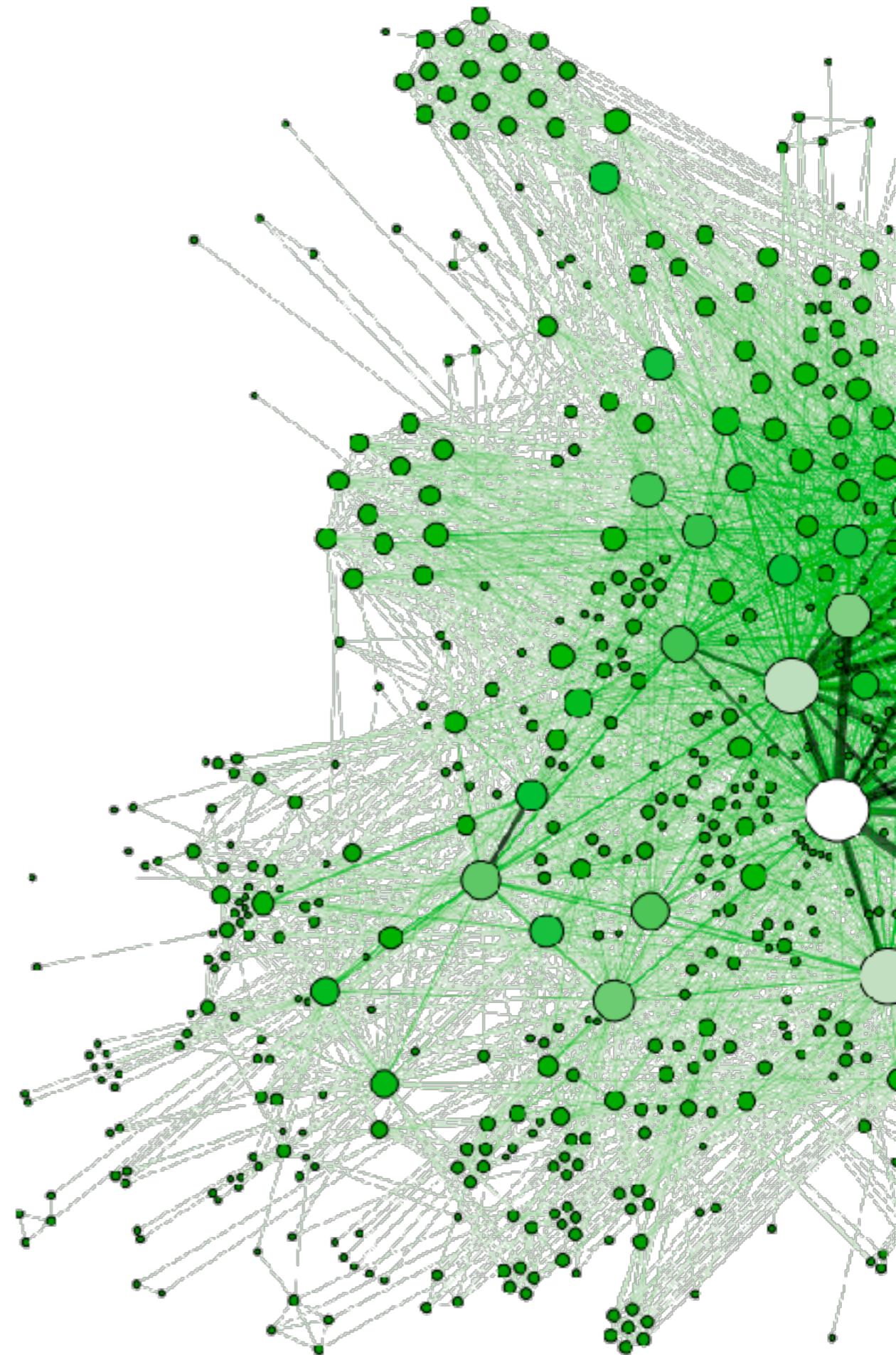
Bachmann, Jennifer; Heimbach, Tabea; Hassenrück, Christiane; Kopprlo, Germa; Iversen, Morten Hvitfeldt; Grossart, Hans-Peter; Gärdes, Astrid (2018): Environmental drivers of free-living vs. particle-attached bacterial community composition in the Mauritania upwelling system. *European Nucleotide Archive (ENA)*, [insdc:PRJEB26997](https://doi.org/10.1093/ena/PRJEB26997)

Project(s):

Leibniz Centre for Tropical Marine Research (ZMT)

- Major upgrade of DMP support tooling (versioning, collaboration, etc.)
- Linking molecular data sets to ORCID profile
- Establishing brokerage for functional genomics data
- Further data repositories for data deposition

NFDI4Biodiversity Service Portfolio



Data submission, versioning and publication

Including long-term preservation
(extended GFBio services)

Helpdesk

Individual support for researchers and
data centers (extended GFBio service)

Support with integration and harmonization of data

(GFBio data centers)

Provision of collaborative workspaces

With support for scientific workflows
and provenance management

Education and Training

Tailored events, tools and materials for
teaching

Basic tools for data managers

Validation, transformation, automated
quality checks various de.NBI tools.

Terminology service

(extended GFBio service)

Search portals and API for data and tools

(extended GFBio services)

Elastic compute service

(Infrastructure-as-a-service)


In future: Research Data Commons

Thank you!

 www.nfdi4biodiversity.org

 www.gfbio.org

 contact@nfdi4biodiversity.org (collaboration)

 info@gfbio.org (service help desk)

 @NFDI4Biodiv



NFDI 4
BIODIVERSITY

Questions?