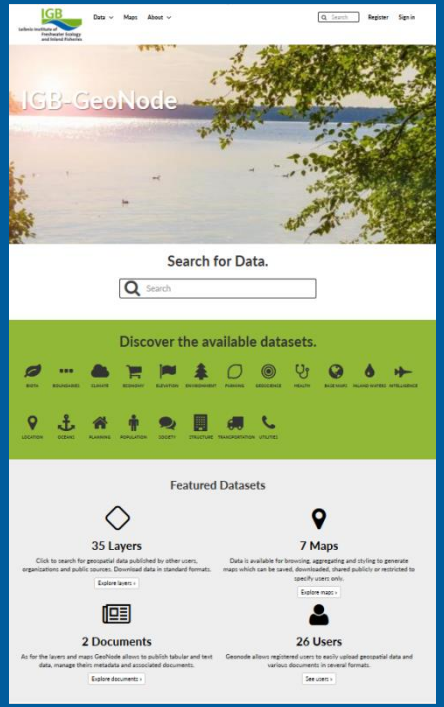
GeoNode Summit

07.12.2020

Leibniz Institute of Freshwater Ecology and Inland Fisheries



FAIR data management from a research institute's perspective:

GeoNode at IGB

Vanessa Bremerich, Sami Domisch, Annett Wetzig, Simone Frenzel, Daniel Langenhaun, Jaime Ricardo Garcia Marquez, Sonja Jähnig

IGB‘s GeoNode task group

Vanessa Bremerich

Sami Domisch

Daniel Langenhaun

Sonja Jähnig

Annett Wetzig

Jaime Ricardo Garcia Marquez

Simone Frenzel

Rita Adrian

IGB in numbers (2019)

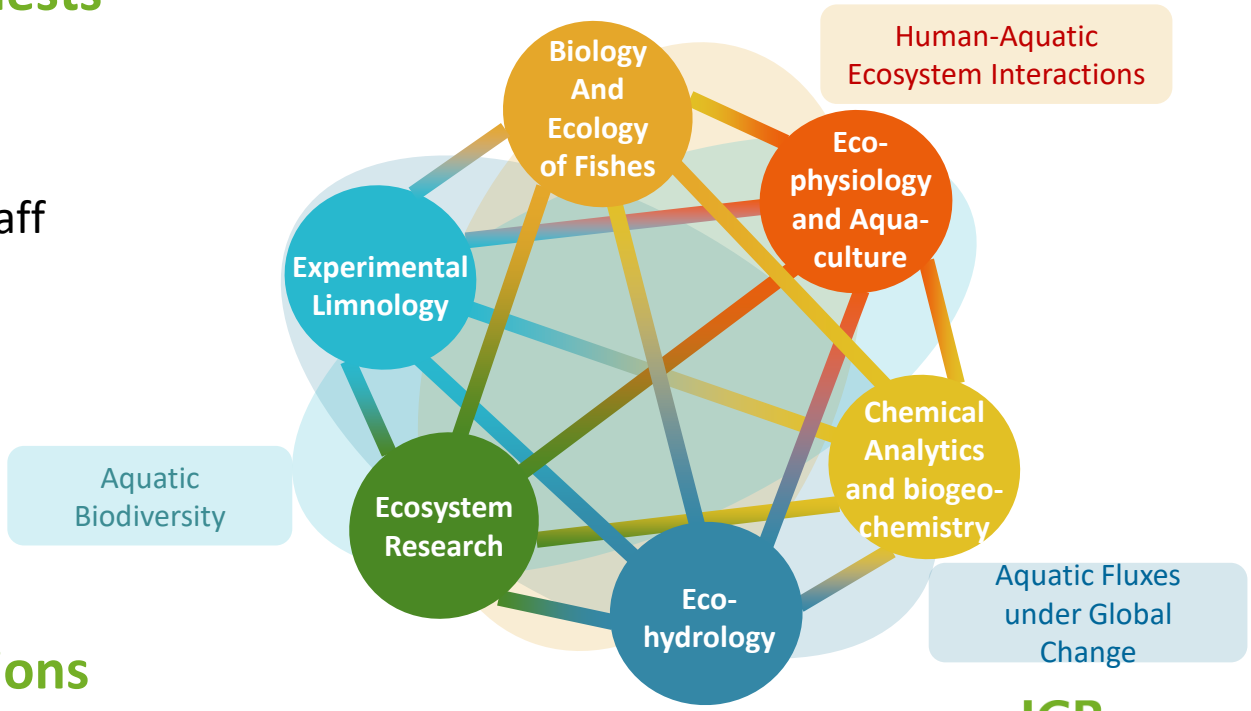
374 employees and guests

- 148 scientists
- 53 doctoral students
- 92 science supporting staff

6 departments

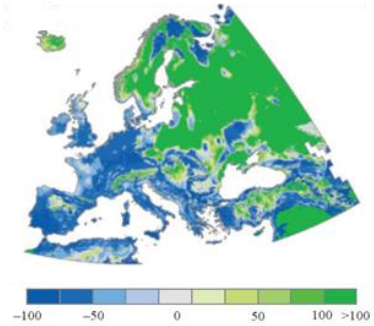
- 38 research groups
- 35 third party funded projects (ongoing)

291 scientific publications



Research data types at IGB

- environmental data, long-term monitoring or sampling data
- biodiversity data
- gene sequencing data
- data from experiments or modeling
- geodata, earth observation data
- social data from questionnaires/surveys
- source code (R, Python,…)



Open Science and the FAIR data principles

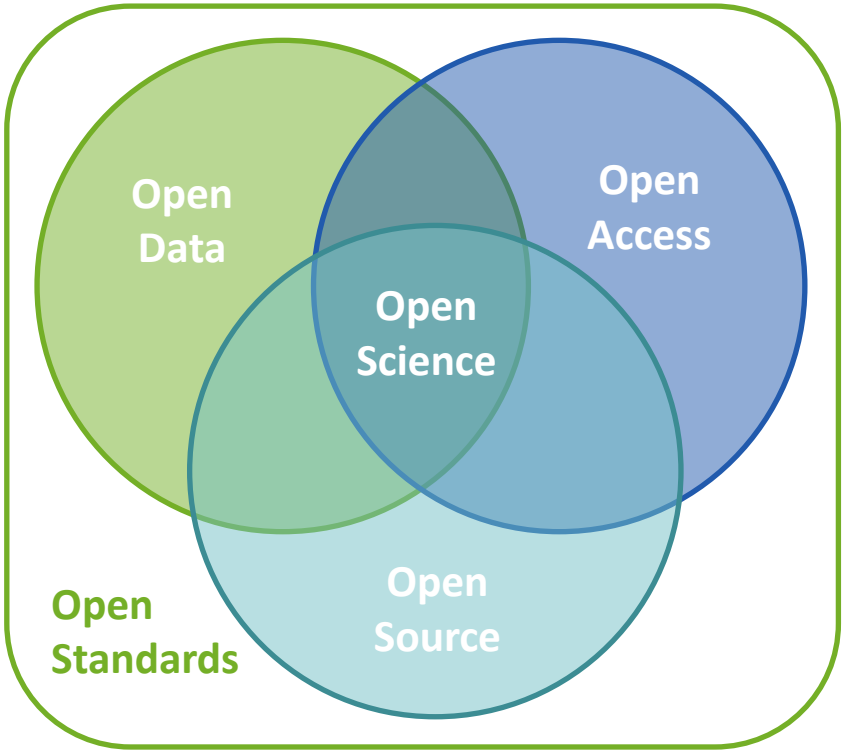


figure altered after Jomier (2017)

F Findable

Data and supplementary materials have sufficiently rich metadata and a unique and persistent identifier.

A Accessible

Metadata and data are understandable to humans and machines. Data is deposited in a trusted repository.

I Interoperable

Metadata and data use a formal, accessible, shared, and broadly applicable language for knowledge representation.

R Reusable

Data and collections have a clear usage licenses and provide accurate information on provenance.

FORCE11, 2014; Wilkinson et al, 2016; LIBER (2017); figure after Goudeseune et al. (2019)

FAIR data at IGB

F

Findable

Data and supplementary materials have sufficiently rich metadata and a unique and persistent identifier.

A

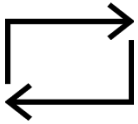
Accessible

Metadata and data are understandable to humans and machines. Data is deposited in a trusted repository.

I

Interoperable

Metadata and data use a formal, accessible, shared, and broadly applicable language for knowledge representation.



R

Reusable

Data and collections have a clear usage licenses and provide accurate information on provenance.

Findable

- Who has what kind of data?
- Are (minimal) metadata added?
- Is a DOI added?

Accessible

- Research data portal available?
- Metadata and data understandable?

Interoperable

- Are standards used (web services, file formats, metadata)?

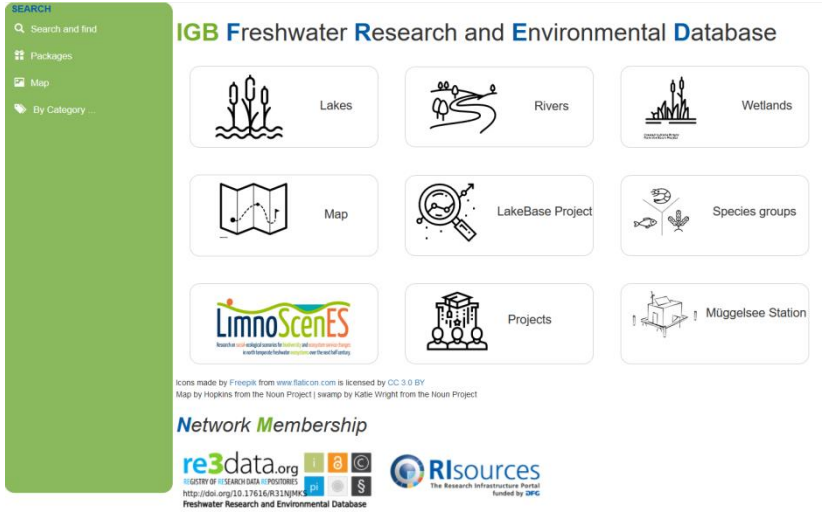
Reusable

- Where is the data from?
- Am I allowed to reuse it?
- What license applies?

Research data management at IGB

FRED - Freshwater Research and Environmental Database

- store and share environmental data and long-term monitoring data
- data from different sources can be combined in „packages“
- mandatory and recommended metadata
- DOI and data license
- semantic search
- data harmonization
- interoperability



https://fred.igb-berlin.de/

GeoNode as IGB‘s geodata management system

Why GeoNode?

- easy-to-use and customizable interface, user-friendly
- metadata automatically added, metadata wizard
- search and filter layers
- user management, access rights and groups
- implementation of OGC standards and protocols

- use of mature open source geospatial projects
- interactive map viewer
- Python/Django framework, relatively easy to extend and build on

→ active community and good documentation, workshops, mailing list, gitter…

GeoNode

IGB GeoNode

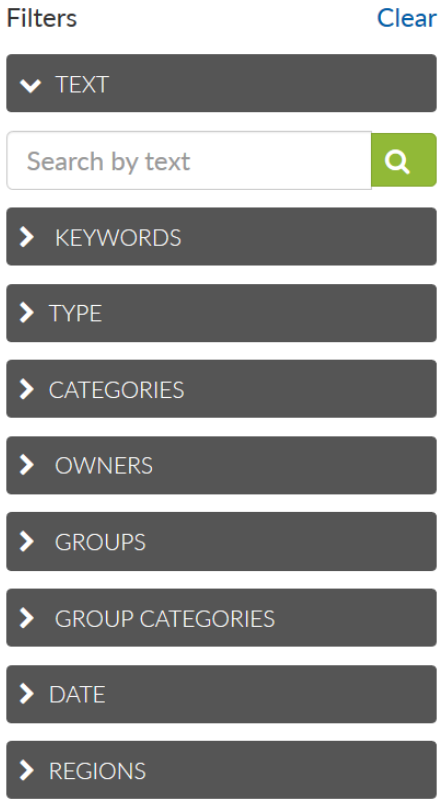
Implementation

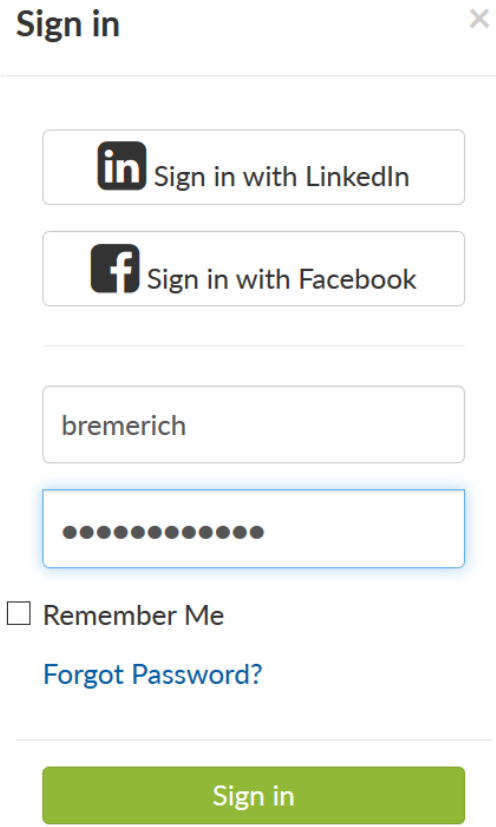
- installation and customization done by GeoSolutions in 2019

→ geo.igb-berlin.de

IGB requirements:

- group and group categories filter
- user authentication via LDAP





→ https://github.com/GeoNode/geonode-contribs/tree/master/ldap

Use cases

1 – Internal geodata management

- examples from current research projects

2 – Link to FRED

- building an integrated research data management infrastructure (FRED and IGB GeoNode) to link geodata to other research data and publications in single "packages"

3 – (Inter)national data and research networks

- Freshwater Information Platform and Global Freshwater Biodiversity Atlas
- extending and building on IGB's GeoNode to develop a pilot for NFDI4Earth

IGB
Leibniz Institute of Freshwater Ecology and Inland Fisheries

IGBFRED

Freshwater Information Platform

Use case 1 – Internal geodata management

GLANCE project

- 5-year research project funded by BMBF
- investigated the impacts of changing flow conditions in rivers on benthic invertebrates and fishes
- three catchments in Germany:
  - Treene in the lowlands,
  - Kinzig in the low mountain ranges
  - Ammer in the Alps

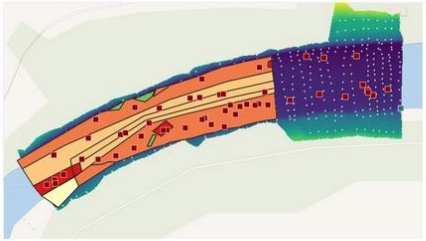
GLANCE

Global change effects in river ecosystems

Sonja Jähnig
Research Group Leader
Research group
Aquatic Ecogeography

https://www.igb-berlin.de/en/project/glance

Use case 1 – Internal geodata management



INLAND WATERS

ABT2

Study site Ammer1 (GLANCE project)

The dataset presented here contains information on the rivers sampled in the GLANCE project. The Ammer1 and 2 catchments are dominated by gravel and stone substrates with steep gradients, which is located in the alpine region of Germany.

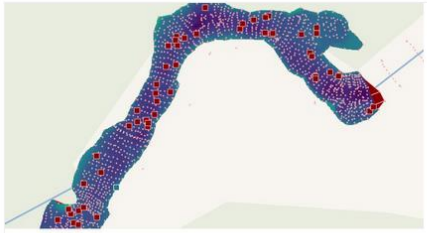
Sonja Jähnig 10 Sep 2020 21 0 0

View Map

Map Layers

This map uses the following layers:

Ammer1_Bathymetry
Ammer_Peissenberg_Bathymetry
Substrates_Ammer1_spring2015
GLANCE SamplingData
SamplingSites_Qmeas



INLAND WATERS

ABT2

Study site Treene (GLANCE project)

The dataset presented here contains information on the rivers sampled in the GLANCE project. The Treene River catchment, with a small gradient, sand-gravel sediment, and groundwater-driven conditions, is located in the northern lowlands of Germany.

Sonja Jähnig 10 Sep 2020 19 0 0

View Map

Map Layers

This map uses the following layers:

Substrates_Treene2_spring2015
Substrates_Treene1_spring2015
Treene1_Bathymetry
Treene_Bathymetry
Treene2_Bathymetry
GLANCE SamplingData
SamplingSites_Qmeas

Use case 1 – Internal geodata management

GLOWABIO
junior research group

- global geospatial analyses
- creating novel high-resolution freshwater ecoregions (90m)
- analyzing spatial patterns in global freshwater biodiversity
- spatial conservation planning given habitat and biodiversity features to highlight potential protection gaps.

GLOWABIO

Global freshwater biodiversity, biogeography & conservation

Sami Domisch
Research Group Leader
Research group
Global Freshwater Biodiversity, Biogeography and Conservation

https://www.igb-berlin.de/en/projekt/glowabio
https://glowabio.org/

Use case 1 – Internal geodata management

GLOWABIO
junior research group

- global scale, high spatial resolution
- all analyses will be performed using open-source geospatial tools (e.g. R, GRASS, GDAL/OGR, pktools, OpenForis)
- all data layers and codes will be stored in public repositories

→ automated pre-processing and upload to GeoNode

GLOWABIO

Global freshwater biodiversity, biogeography & conservation

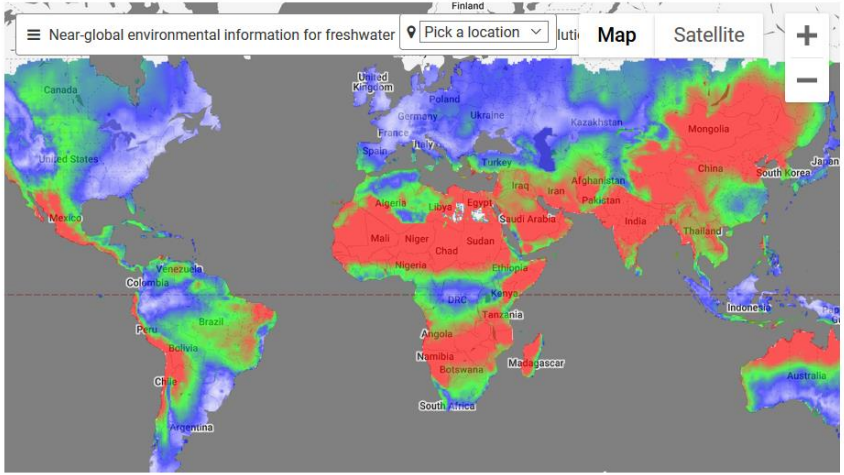
Sami Domisch
Research Group Leader
Research group
Global Freshwater Biodiversity, Biogeography and Conservation

https://www.igb-berlin.de/en/projekt/glowabio
https://glowabio.org/

Use case 1 – Internal geodata management

Freshwater-specific environmental variables at 1km resolution

Near-global freshwater-specific variables at 1km spatial resolution (Domisch et al. 2015). Data download at earthenv.org/streams.



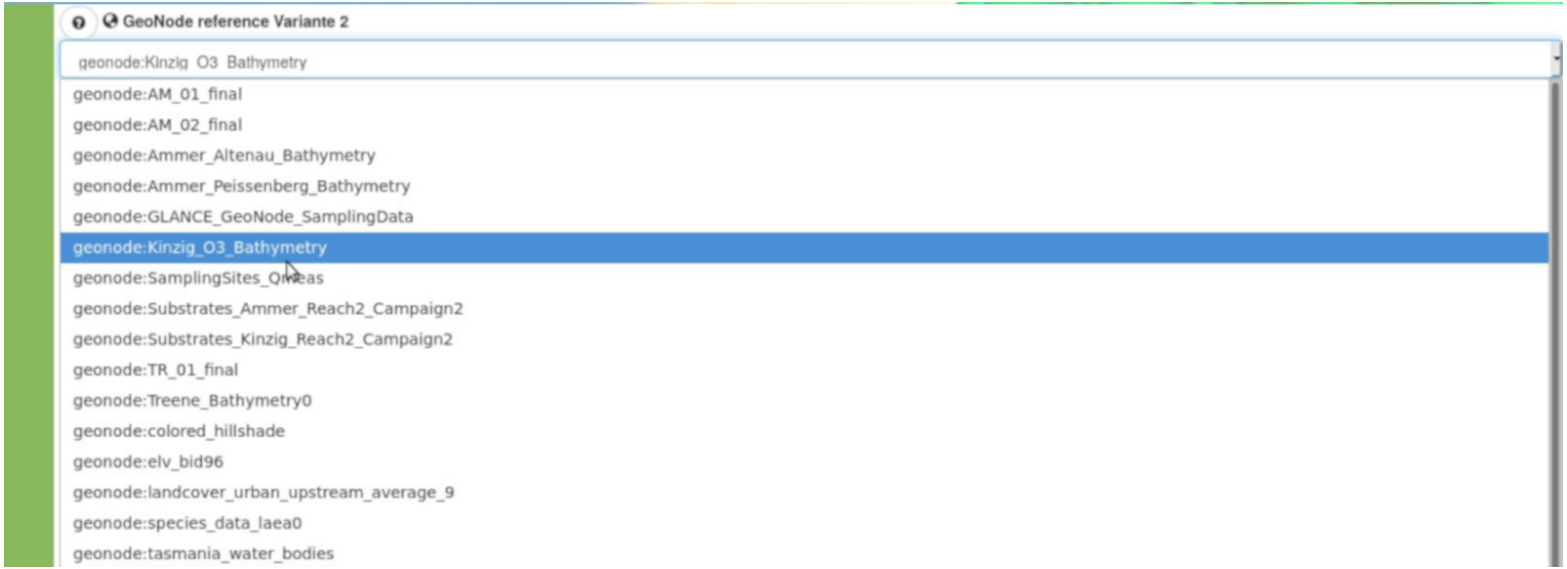
https://glowabio.org/maps/1km_variables/

Test data for freshwater distribution models

For illustration purpose: we use example stream networks, such as the one below that can be quickly generated in GRASS-GIS, for testing various tools and techniques for routing, network modeling, point processing etc.

https://glowabio.org/maps/sdm_test_data/

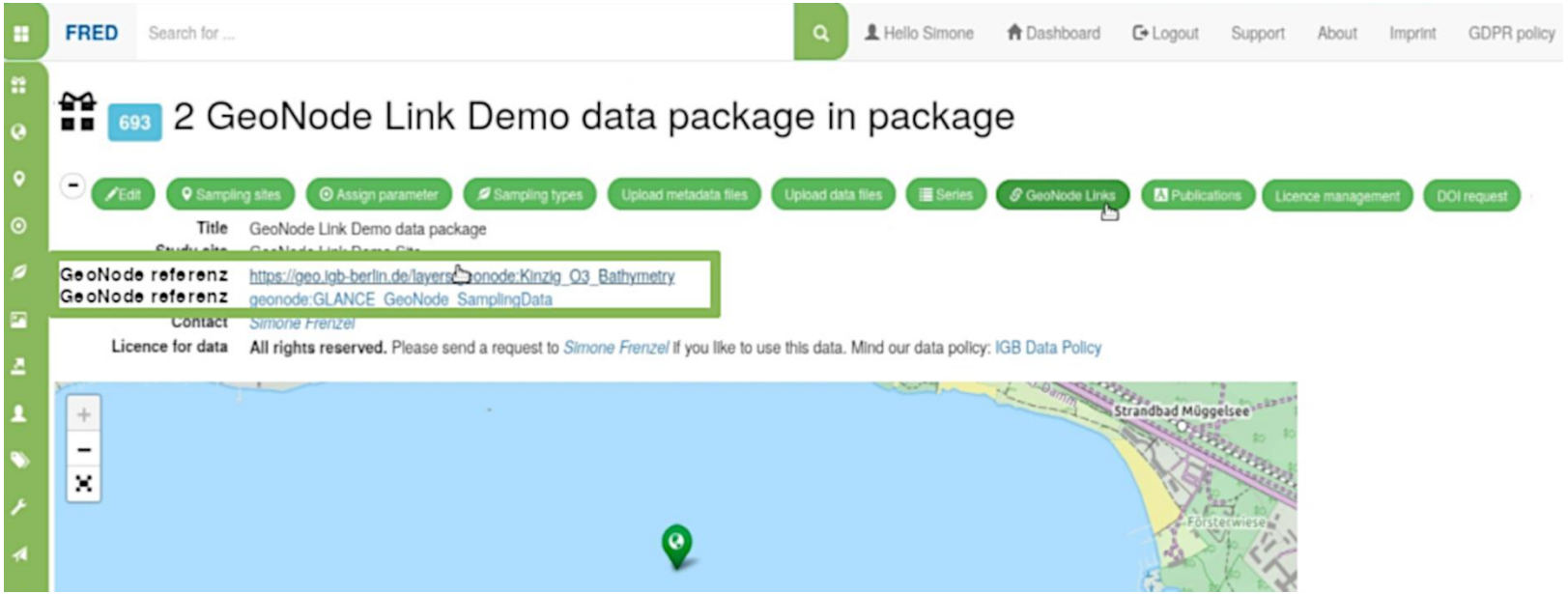
Use case 2 – link to FRED



GeoNode to FRED:

- layer list from GeoNode read via REST

Use case 2 – link to FRED



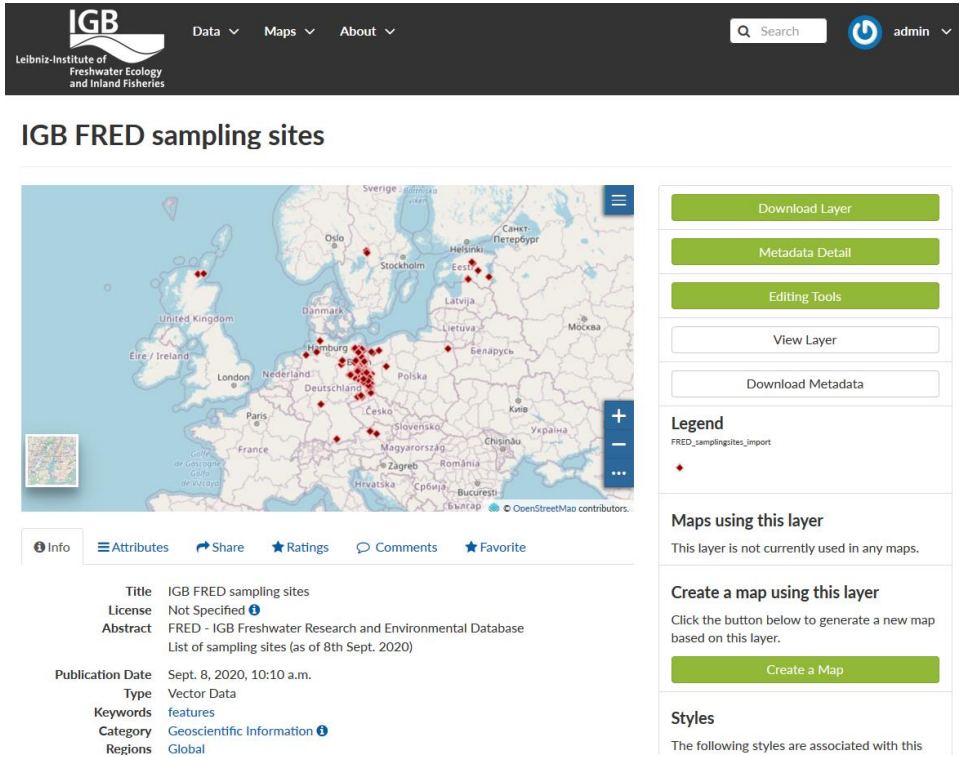
GeoNode to FRED:

- selected layer(s) can be added to data package

Use case 2 – link to FRED

FRED to GeoNode

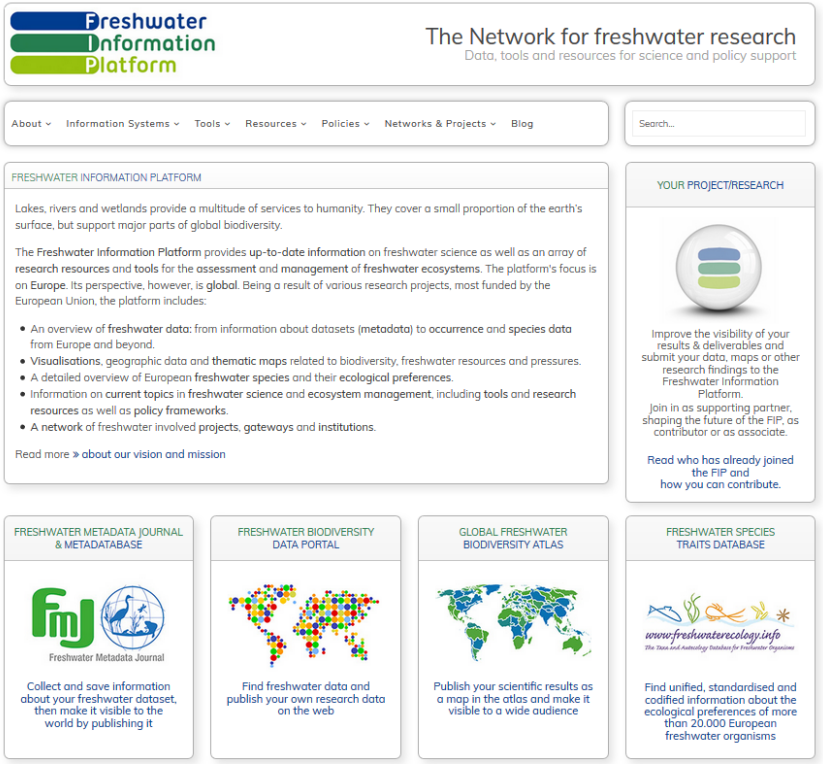
- read access on selected PostgreSQL views to evaluate direct import into GeoServer
- e.g. sampling sites, study sites, long-term monitoring data



Use case 3 – (Inter)national data and research networks

IGB GeoNode as geospatial repository

- for the Freshwater Information Platform and the
- Global Freshwater Biodiversity Atlas



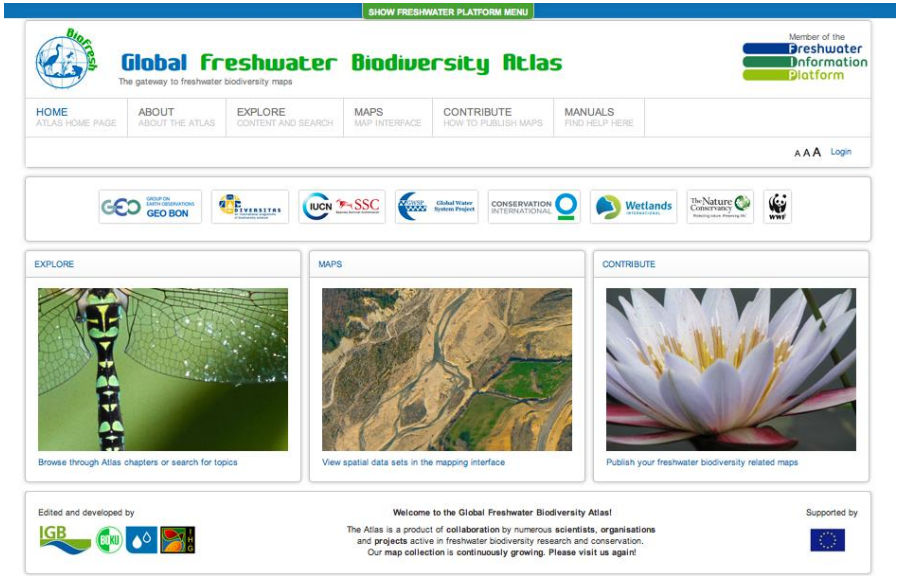
Freshwater Information Platform

http://www.freshwaterplatform.eu/

Use case 3 – (Inter)national data and research networks

Global Freshwater Biodiversity Atlas

- established in 2013
- EU FP7 project BioFresh
- focus on broad scale freshwater biodiversity maps



Freshwater Information Platform

http://atlas.freshwaterbiodiversity.eu /

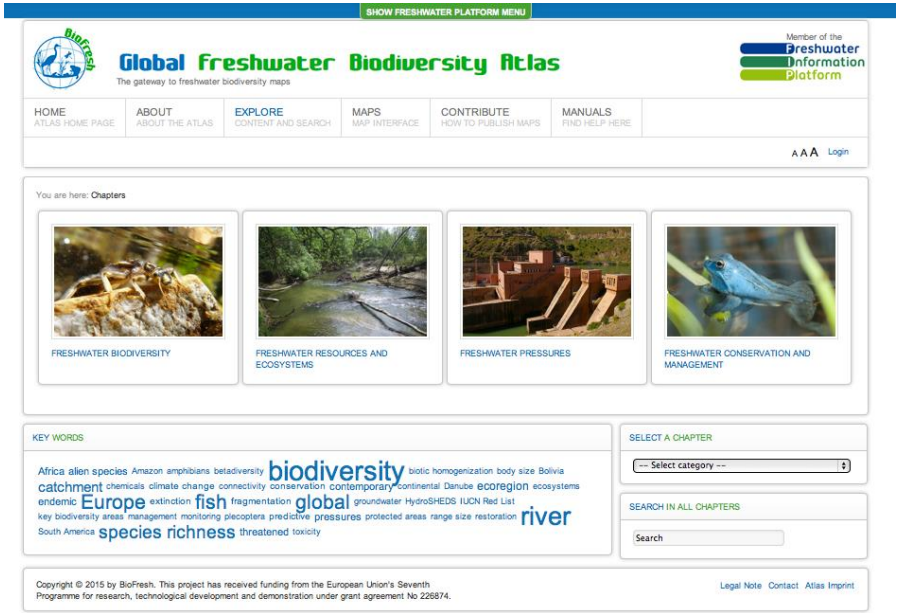
Use case 3 – (Inter)national data and research networks

Global Freshwater Biodiversity Atlas

maps in four chapters:

- Freshwater Biodiversity
- Resources and Ecosystems
- Freshwater Pressures
- Conservation and Management

Freshwater Information Platform



http://atlas.freshwaterbiodiversity.eu/

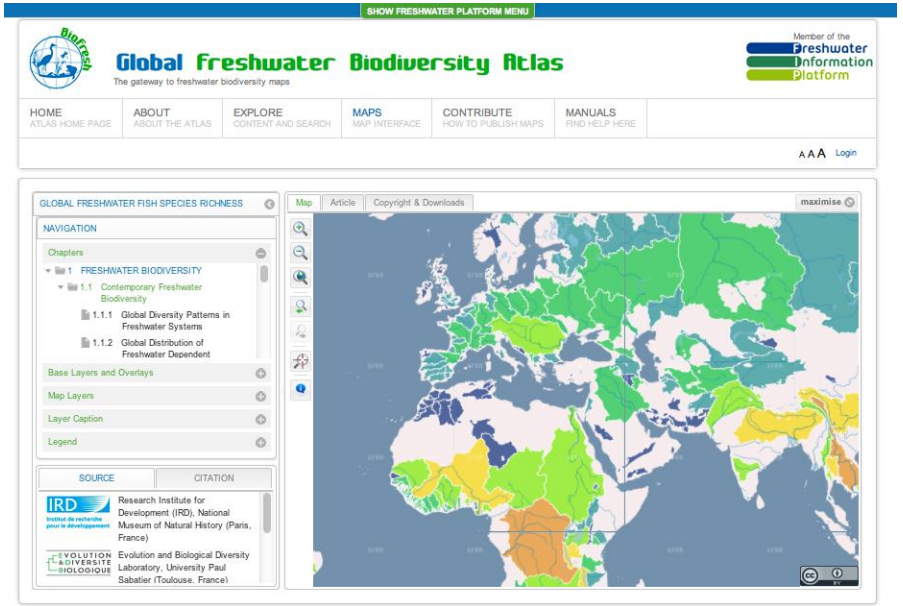
Use case 3 – (Inter)national data and research networks

Global Freshwater Biodiversity Atlas

- interactive GeoExt2 mapping tool
- GeoServer

→ migrate to GeoNode

→ metadata and download links



Freshwater Information Platform

http://atlas.freshwaterbiodiversity.eu/

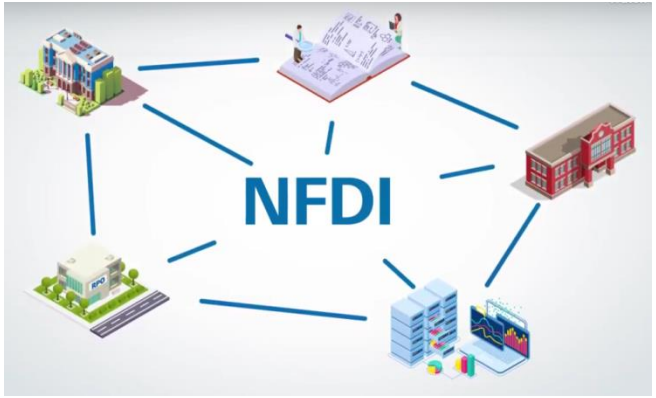
National research data infrastructure (NFDI)

NFDI

is a coordinated network of consortia tasked with providing science-driven data services to research communities.

Aim

- systematically manage scientific and research data
- provide long-term data storage, backup and accessibility
- network the data both nationally and internationally



https://www.dfg.de/en/research_funding/programmes/nfdi/index.html

https://youtu.be/x3Cvn1vNQ98

NFDI$_4$Earth

National Research Data Infrastructure for Earth System Science (NFDI4Earth)

- funding decision of the Joint Science Conference (GWK) in June 2021
- five years' work plan (2021-26)
- currently 54 partners

Aim

- provide researchers with FAIR, easy, coherent, efficient (and open and unrestricted) access to all relevant Earth System data, scientific data management tools and data analysis services.

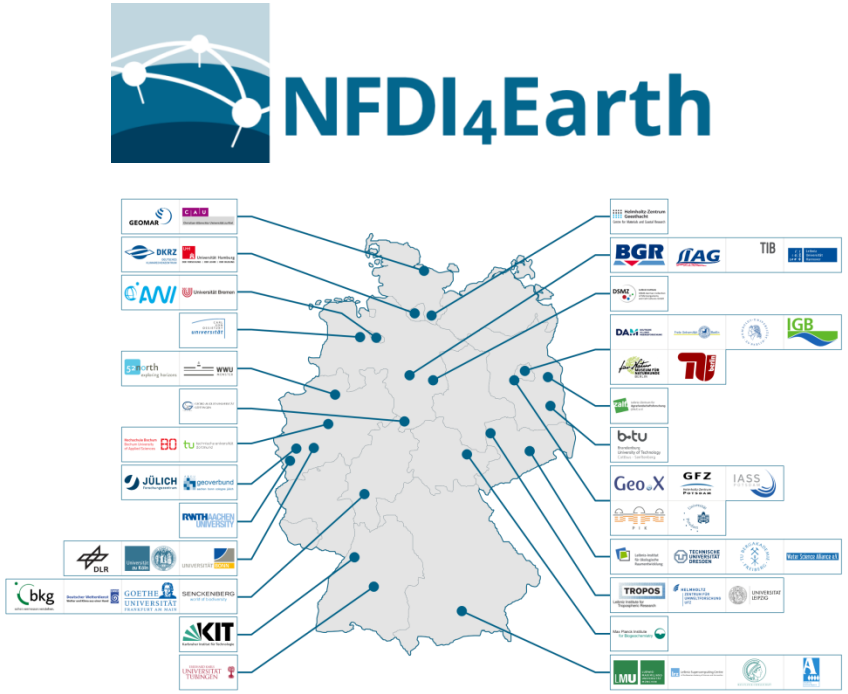


figure: https://www.nfdi4earth.de/about-us/ambition-and-mission

NFDI$_4$Earth Pilots 2020

Search

Topics

- 3 Atmospheric Science, Oceanography and Climate Research
- 2 Geography
- 2 Geology and Paleontology
- 2 Geophysics and Geodesy

Pilot Principial Investigator

- 2 Deutsches Zentrum für Luft- und Raumfahrt
- 2 Technische Universität Dresden
- 1 Deutscher Wetterdienst
- 1 GEOMAR Helmholtz-Zentrum für Ozeanforschung Kiel

Type of Institution

- 6 University
- 4 Helmholtz
- 2 Leibniz
- 1 Departmental Research
- 1 Max-Planck

14 Items filtered from 38 originally (*Reset All Filters*)

Name	Topics	Suggested for funding	Proposal
Bathy4All: Workflows for Multibeam Processing and Visualization	Geophysics and Geodesy	after 1st call	PDF
Enhancing Earth system model evaluation with data cube enabled machine learning	Atmospheric Science, Oceanography and Climate Research	after 1st call	PDF
GeoRoCc-MetBase Tools	Geochemistry, Mineralogy and Crystallography	after 1st call	PDF
German marine seismic data access	Geology and Paleontology	after 1st call	PDF
Getting freshwater spatio temporal data on track (GeoFRESH)	Water Research	after 1st call	PDF
Interoperability and reusability for geoscientific lab data	Geophysics and Geodesy	after 1st call	PDF
Linking Environmental D			

https://https://www.nfdi4earth.de/participate/get-involved-by-pilots

NFDI$_4$Earth Pilots 2020 - GeoFRESH

NFDI$_4$Earth

Getting freshwater spatiotemporal data on track

Sami Domisch[1],Giuseppe Amatulli[2], Vanessa Bremerich[1], Luc De Meester[1], Mark Gessner[1], Hans-Peter Grossart[1], Rita Adrian[1]

[1]Leibniz-Institute of Freshwater Ecology and Inland Fisheries (IGB), Müggelseedamm 301, 12587 Berlin, Germany
[2]Yale University, Centre for Research Computing, New Haven, CT, 06511, USA

Date of submission: July 24th 2020

Abstract

Spatiotemporal freshwater-related earth system data are currently poorly organized and its full potential for research or management is rarely exploited, due to data disparity and its missing interoperability given the different data standards and formats. It is especially the spatial struc-

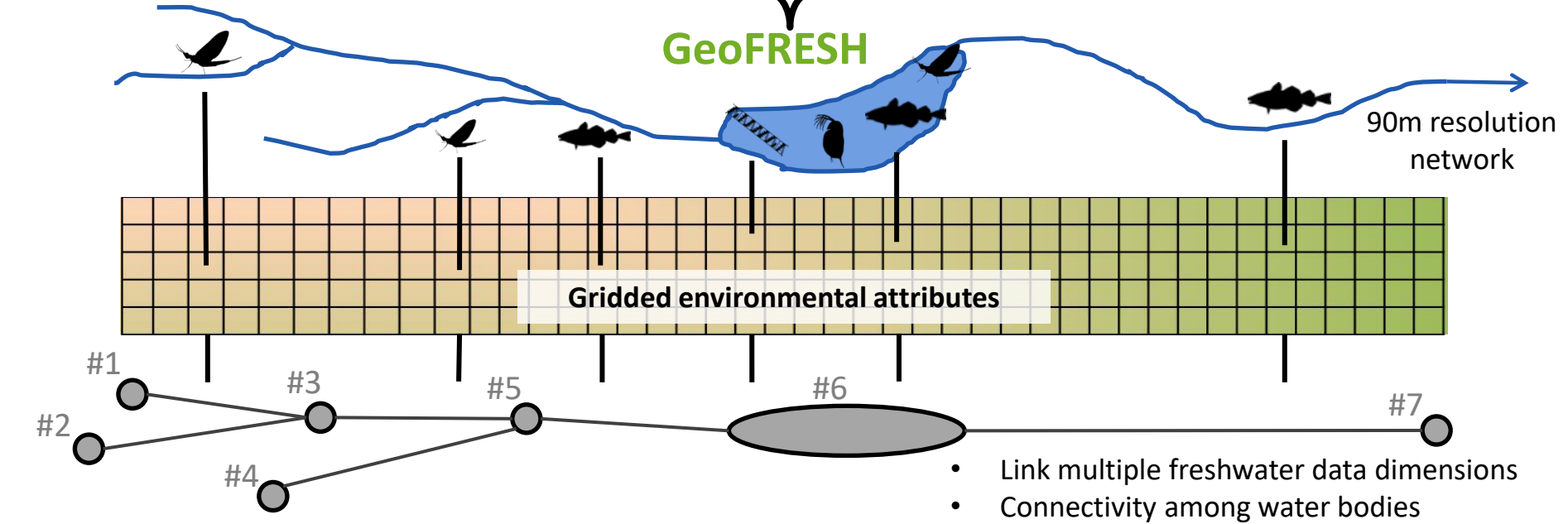
https://nfdi4earth.de/images/nfdi4earth/documents/pilots/proposals/116-Getting_freshwater_spatiotemporal_data_on_track_GeoFRESH.pdf

Freshwater spatiotemporal data integration within an online platform

FRED ⟷ IGB GeoNode

Survey data | Citizen science | Museum collections | Opportunistic data | Gridded data | Polygon data

GeoFRESH



- Link multiple freshwater data dimensions
- Connectivity among water bodies
- FAIR data for Earth system sciences

- Jomier (2017). Open Science – Towards Reproducible Research. 361 - 367.
- FORCE11 (2014). The Fair Data Principles. Retrieved from: https://www.force11.org/group/fairgroup/fairprinciples
- Wilkinson M. D., Dumontier M., Aalbersberg, I. J., Appleton G., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3, 160018. http://doi.org/10.1038/sdata.2016.18
- LIBER (2017). Implementing FAIR Data Principles: The Role of Libraries. LIBER (Association of European Research Libraries) factsheet. https://libereurope.eu/wp-content/uploads/2020/09/LIBER-FAIR-Data.pdf
- Goudeseune et al. (2019). Guidance document for scientists on data management, open data, and the production of Data Management Plans. BiodivERsA report. 48 pp.