

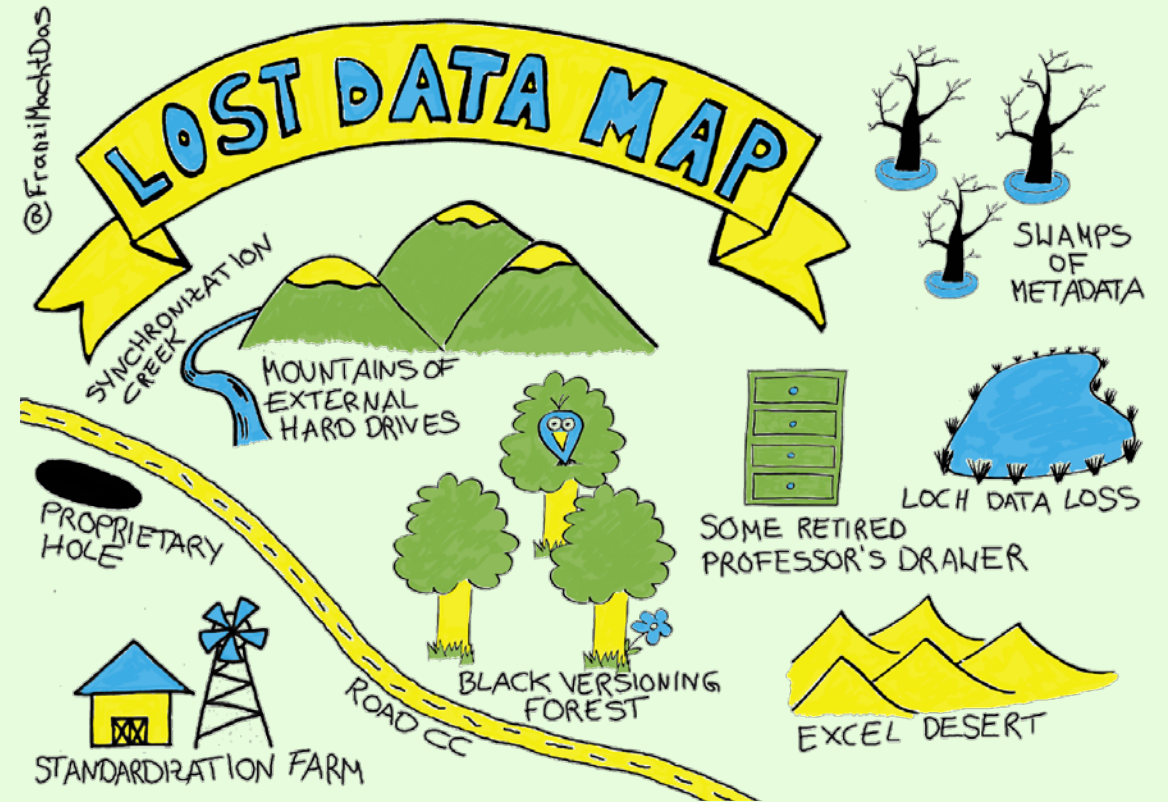
NFDI4Biodiversity – Biodiversity, Ecology and Environmental Data

[FAIR data!]



Our Goal: Making Research Data FAIR

F	FINDABLE
A	ACCESSIBLE
I	INTEROPERABLE
R	REUSABLE



Services in NFDI4Biodiversity

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**

Data types and archives

Typical types of data our partners can handle:

- Occurrence Data (species)
- Environmental data (e.g. temperature, rainfall)
- Trait data (e.g. seed number and mass)
- Molecular data (e.g. sequences)
- Experimental and laboratory measurements
- Multimedia (photographs, audio, video), f.e. of observed specimens
- Orthophotos produced using a drone
- Digital surface models
- Model code
- Statistics

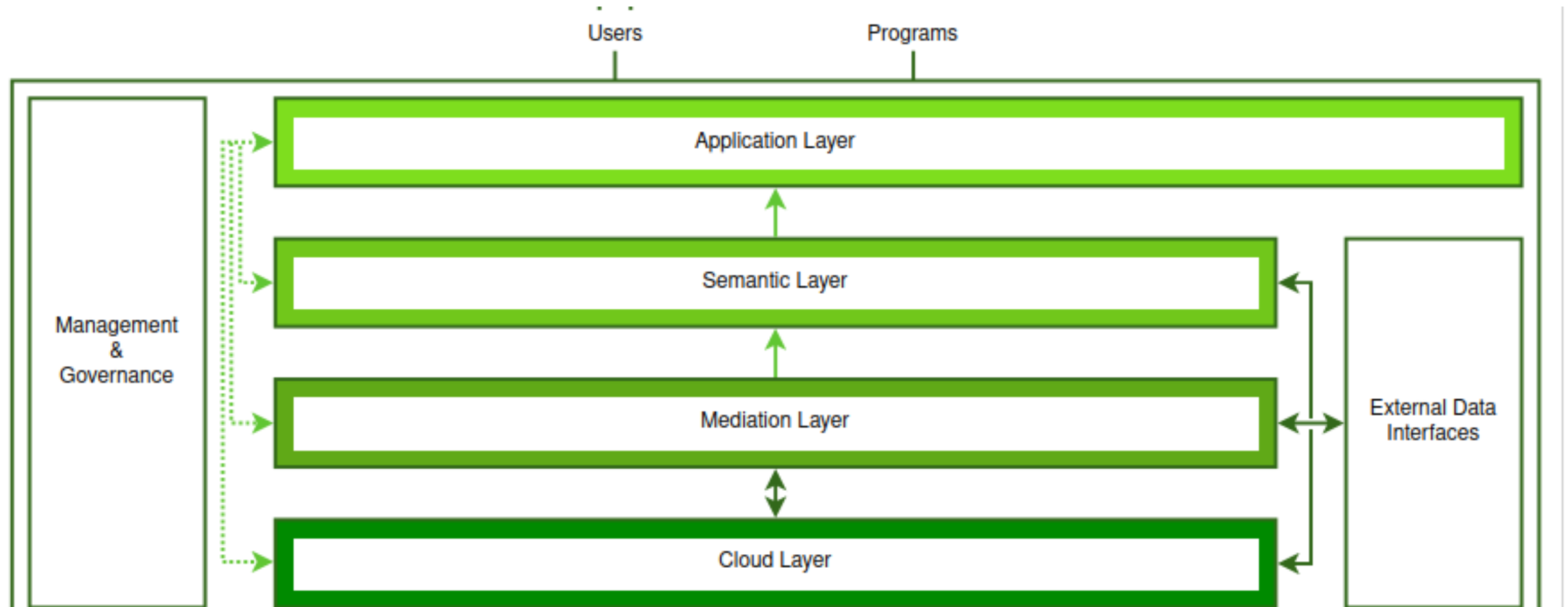
Data Centers specialized on Nucleotide, Plant and Environmental Data



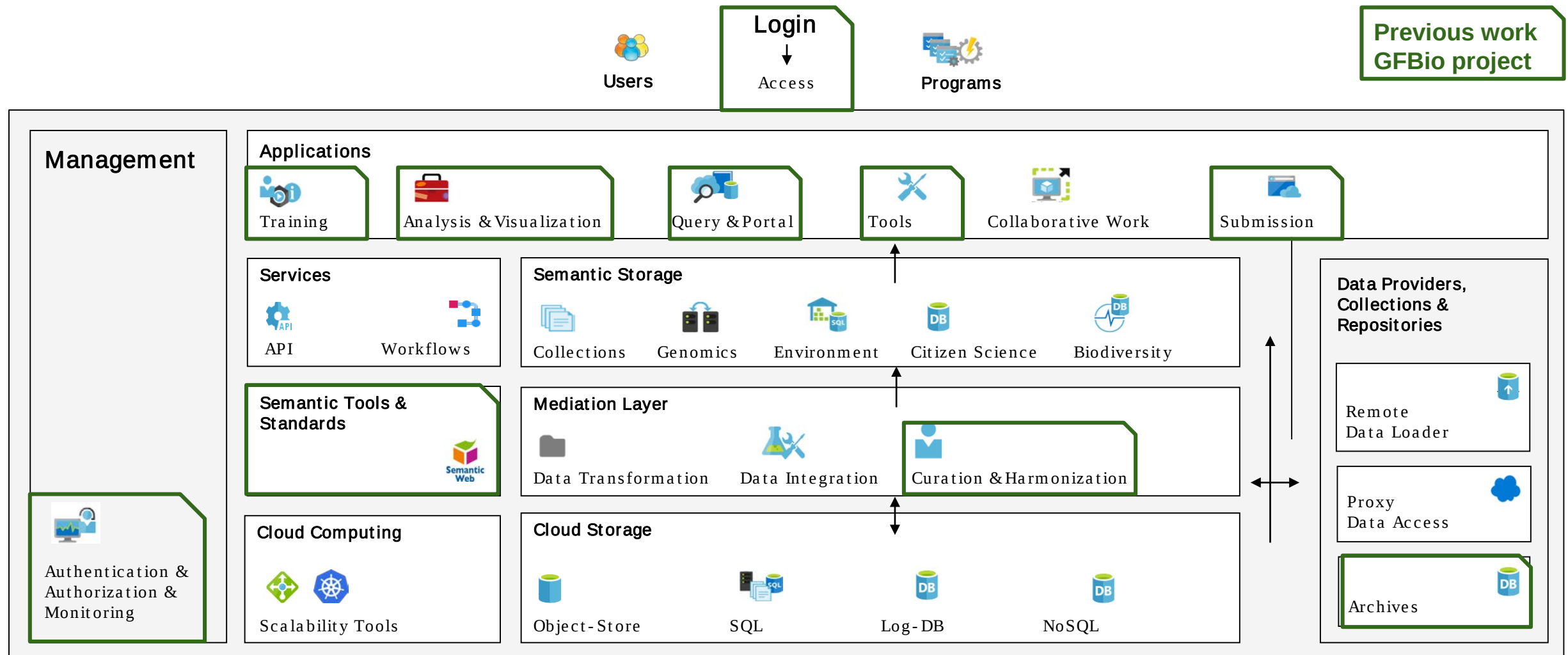
Data Centers at Natural Science Collections



Moving services to the academic cloud



Research Data Commons – available components

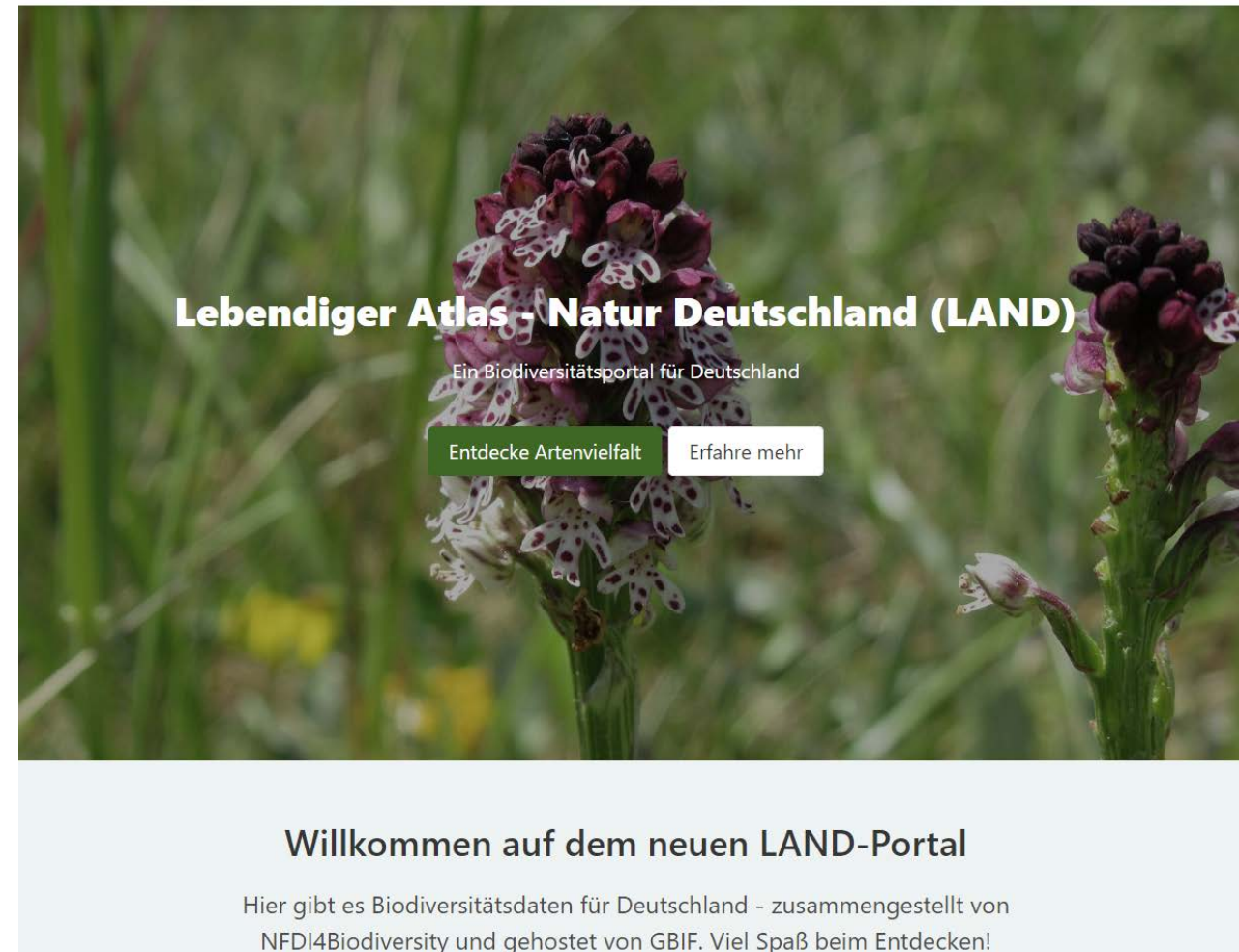


Selected cooperations:

GBIF (Global Biodiversity Information Facility)

- Setup of a GBIF hosted portal for observation data ⇒ “Lebendiger Atlas Natur Deutschland (LAND)”
<https://land.gbif.de/>
- Moderated by the BioCASE helpdesk @ FU Berlin-BO
- Provides a standardized data workflow, DQ components, and a capable data portal
- Establishes a GBIF-pipeline and cooperation based on a concrete technical product

<https://land.gbif.de/>



GBIF... continued

Issue: Reference for species; persistent identifier

Relevant for archiving, tracking, synonyms, changed taxonomy



Sources for species IDs

- GBIF key (example: [Bombus terrestris](https://www.gbif.org/species/xxx)) => <https://www.gbif.org/species/xxx>
- Catalogue of Life ID (CoL), e.g. [Bombus terrestris](https://www.catalogueoflife.org/data/taxon/xxx) => <https://www.catalogueoflife.org/data/taxon/xxx>
- Integrated Taxonomic Information System ITIS ID (<https://itis.gov/>), [Bombus terrestris](https://itis.gov/)

Tools

- GBIF Species matching (<https://www.gbif.org/tools/species-lookup>)
- CoL & GBIF: Checklistbank (<https://www.checklistbank.org/tools/name-match>)
- ITIS: Compare Taxonomy/Nomenclature (<https://www.itis.gov/tools.html>)

Selected cooperations:

Machine-to-machine data exchange

Access to Checklist data

- Development of API - additional programming funded through NFDI4Biodiversity project budget
- Formalized agreement between GFBio e.V. (NFDI4Biodiversity development partner) and BfN (contracting authority Checklist Editor)



NFDI 4
BIODIVERSITY

Selected cooperations:

Institutional Data Management

- Hosted Data Management Information System (**BEXIS**); service provided for **Nationalpark Bayerischer Wald** by NFDI4Biodiversity partner **University Jena**
- Prototyping archival workflows with Bavarian State Archive (GdA; **Staatliche Archive Bayerns**)
- Cooperation governed within the Consortium (all three are partners)



Selected cooperations:

Nationale Naturlandschaften (NNL)

Requirements for many NNL members

- Cloud storage
- Analysis tools



NFDI4Bio Network: options for solutions

⇒ GWDG (Gesellschaft für
Wissenschaftliche Datenverarbeitung
mbH Göttingen; Max Planck, Uni
Göttingen; <https://gwdg.de>)



Selected cooperations:

Data driven methodology



IUCN Red list fish species

- For the first time a data driven approach for short term trends
- > 6 month to collect and harmonize the data
- Data set will be archived and made publicly available
- Accompanying data publication in preparation



NFDI 4
BIODIVERSITY

Selected cooperations:

Project initiated data pipeline



Data provision for eLTER IT infrastructure:

1. Hosted Data Management Information System **BEXIS** for LTER- D; data structure and metadata scheme adjusted to eLTER SO (data base, quality check; BEXIS Service provided by NFDI4Biodiversity partner University Jena)
2. (semi-) automated export to **Pangaea**
3. Pangaea: DOI, data publication (landing page)
4. eLTER data: harvested from Pangaea by eLTER IT infrastructure



Selected cooperations:

Project initiated data pipeline



Actual status

- eLTER use case: workflow elements to be implemented
 - Bexis 4 (will be released in March 2025)
 - eLTER Metadata schema (json) for eLTER Standard Observations ready for implementation in Bexis
 - Species name check (to be decided how this will be done)



Resources on Youtube



Data Management for Researchers [GFBio]

5 Videos • 21 Aufrufe • Zuletzt am 10.11.2021 aktualisiert

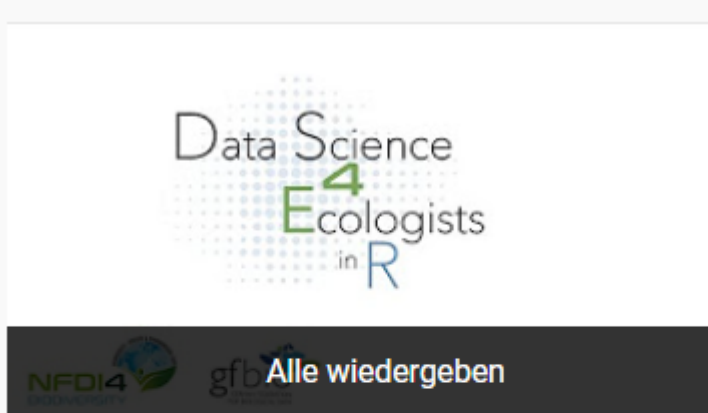
≡+ ↺ ↻ ...



Data Visualization, Analysis and Transformation

5 Videos • 38 Aufrufe • Zuletzt am 07.12.2021 aktualisiert

≡+ ↺ ↻ ...



Data Science 4 Ecologists in R

3 Videos • 96 Aufrufe • Zuletzt am 07.12.2021 aktualisiert

≡+ ↺ ↻ ...

Thank you!



NFDI 4 BIODIVERSITY



U N I K A S S E L
V E R S I T Ä T



Funded by the DFG, project number 442032008

18.03.20

25

Titel des Meetings / der Präsentation

17