

Die Rolle der DWB in der Daten- und Sammlungsinfrastruktur des LIB

Birgit Rach
LIB Museum Koenig, Bonn

Das LIB - Leibniz-Institut zur Analyse des Biodiversitätswandels

- Forschungsmuseum der Leibniz-Gesellschaft
- An zwei Standorten: Museum Koenig, Bonn und Museum der Natur, Hamburg
- Diverse Abteilungen:
 - Verschiedene Zoologische Sammlungen
 - Geologische Sammlung
 - Paläontologische Sammlung
 - Biobank (Gewebe, DNA, Zellen in Kryolagerung)
 - Umweltprobenbank (im Aufbau, u.a. Insektenmischproben)
 - Biohistoricum
- Aktuelle Forschungsprojekte
- GFBio Datencenter



Komponenten der Daten- und Sammlungsinfrastruktur am LIB

- **DWB** – Framework für Sammlungsdaten und Metadaten



- **fyI_r** – Asset Management System für Bilder und Filme (<https://media.leibniz-lib.de>)



- **Geneious** – Software zur Analyse von Sequenzdaten



- **ASV-Registry** – Plattform zum Managen von ASV- und OTU-Tabellen

(<https://data.bolgermany.de/gbol1/metabarcoding>)

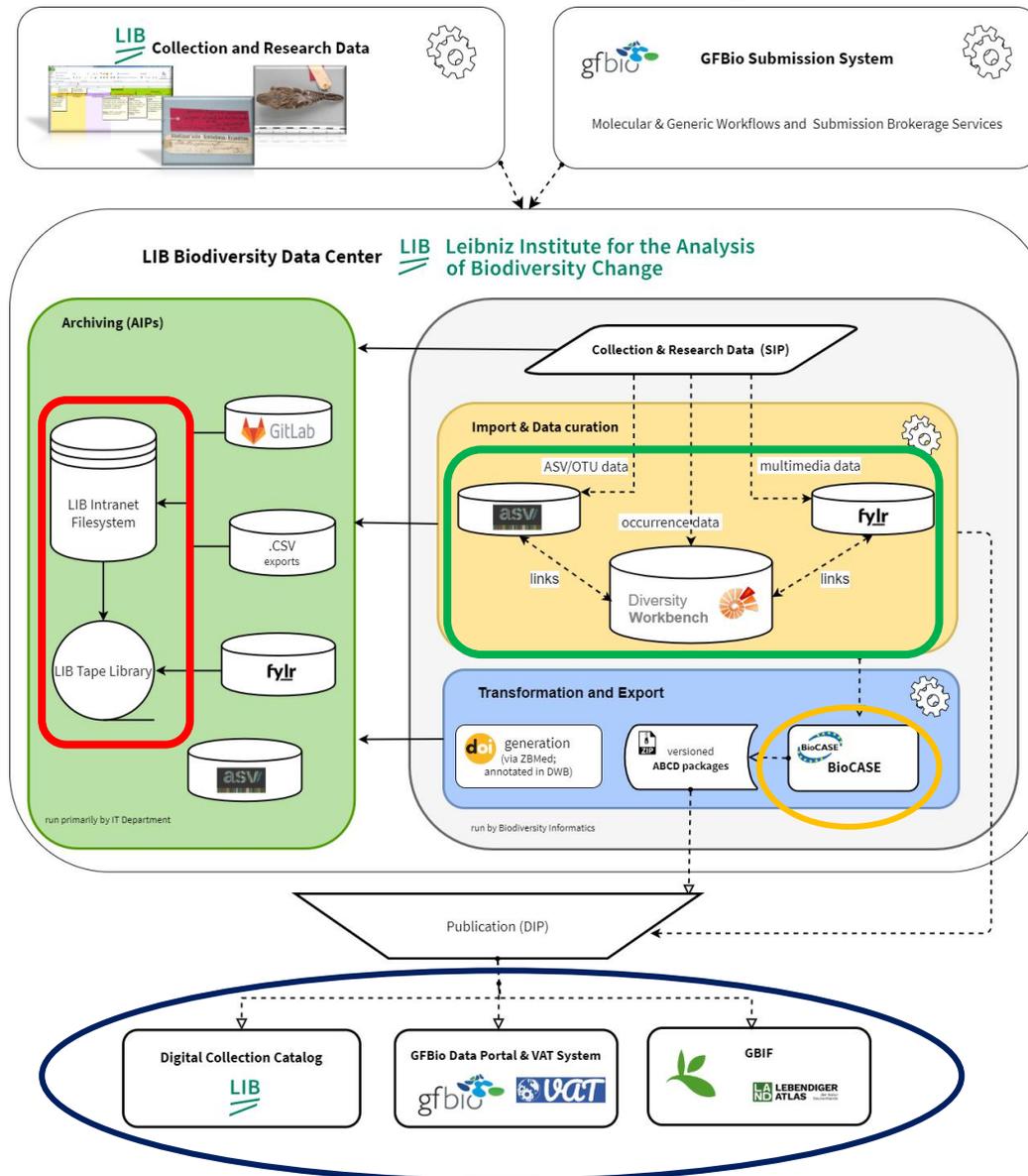


- **BioCASE Provider Software** - Publishing Software (z.B. zu GBIF, GfBio /NFDI)



- **Digitaler Sammlungskatalog** (<https://collections.leibniz-lib.de>)

Komponenten der Daten- und Sammlungsinfrastruktur am LIB



Sammlungsdatenverwaltung



Import der Sammlungsdaten mit dem Import Wizard

- Sammlungsspezifische Importschemata
- Abgleich zwischen den Sammlungen BN/HH (Workflows, external identifier, Analysen, etc.)

Eintrag einzelner Datensätze

File: C:\Users\brach\OneDrive\GFBO\Projekte\DSUB-558-Metabarcoding\Arthropods SW Germany\2_Processing\Importe\Ausgangdaten\Ausgangdaten_DNA_MB_2021_mt_ASV-Cv.txt

Encoding: UTF-8

Separator: TAB Language / Country: US

Start line: 3 End line: 2803

Translate \r\n to line break

Record all SQL statements

Win to return: Encoding Unicode Separator: tab Language: US

First line contains column definition: Use default duplicate check:

Attached to table: CollectionSpecimen Column: AccessionNumber

Schedule for import of tab-separated text files into DiversityCollection

Target within DiversityCollection: Specimen

Versions: Software 4.4.15 Database 02.06.48 Schedule 1 Lines 3 - 2803

Columns in table

Column in table	Key	Copy	Pre	Post	File	pos.	transformations	Value	Source	Table
CollectionEventID									Database	
CountryCache					52	Filter	Import fixed value: Germany if column 52 = 0		File	

CollectionSpecimen (CollectionSpecimen)

Record identifier	Record URL	Date last edited	RECORDBASIS	BOLD_Process_ID	BOLD_BBN_Lat	BOLD_Grade/UID	BOLD_M_Seq_L	adjusted_Phylum	adjusted_Class	adjusted_Order	adjusted_Family	adjusted_Genus	adjusted_Species
AAATAAATGTT.P142890-1	BOLD link as file	last update of ref	Kind of record: Lab	Internal sample n.	Name of the file:	Overlap between:	Number of identic.	Name of the phyl.	Name of the class.	Name of the order.	Name of the family.	Name of the genus.	Name of the species.
AAATAAATGTT.P142890-2	http://www.bolds...	05.11.2024	MachnoOceana	NLEA888-12	BOLD-AAA1236	96.99%	313	Arthropoda	Insecta	Lepidoptera	Hemiptera	Lepidoptera	Lepidoptera
AAATAAATGTT.P142890-3	http://www.bolds...	05.11.2024	MachnoOceana	ROBA638-06	BOLD-AAA1471	100.0%	313	Arthropoda	Insecta	Lepidoptera	Hemiptera	Lepidoptera	Lepidoptera
AAATAAATGTT.P142890-4	http://www.bolds...	14.10.2024	MachnoOceana	PHLAH721-12	BOLD-AAA1496	100.0%	313	Arthropoda	Insecta	Lepidoptera	Hemiptera	Lepidoptera	Lepidoptera
AAATAAATGTT.P142890-5	http://www.bolds...	14.10.2024	MachnoOceana	PHLAH721-12	BOLD-AAA1513	100.0%	313	Arthropoda	Insecta	Lepidoptera	Hemiptera	Lepidoptera	Lepidoptera
AAATAAATGTT.P142890-6	http://www.bolds...	14.10.2024	MachnoOceana	NLEA1111-12	BOLD-AAA2052	96.99%	313	Arthropoda	Insecta	Lepidoptera	Hemiptera	Lepidoptera	Lepidoptera
AAATAAATGTT.P142890-7	http://www.bolds...	14.10.2024	MachnoOceana	PHLAG947-12	BOLD-AAA2076	96.99%	313	Arthropoda	Insecta	Lepidoptera	Hemiptera	Lepidoptera	Lepidoptera
AAATAAATGTT.P142890-8	http://www.bolds...	14.10.2024	MachnoOceana	PHLAG947-12	BOLD-AAA2144	96.99%	313	Arthropoda	Insecta	Lepidoptera	Hemiptera	Lepidoptera	Lepidoptera
AAATAAATGTT.P142890-9	http://www.bolds...	14.10.2024	MachnoOceana	PHLAG947-12	BOLD-AAA2177	96.99%	313	Arthropoda	Insecta	Lepidoptera	Hemiptera	Lepidoptera	Lepidoptera
AAATAAATGTT.P142890-10	http://www.bolds...	14.10.2024	MachnoOceana	PHLAG947-12	BOLD-AAA2224	96.99%	313	Arthropoda	Insecta	Lepidoptera	Hemiptera	Lepidoptera	Lepidoptera
AAATAAATGTT.P142890-11	http://www.bolds...	14.10.2024	MachnoOceana	PHLAG947-12	BOLD-AAA2224	100.0%	313	Arthropoda	Insecta	Lepidoptera	Hemiptera	Lepidoptera	Lepidoptera
AAATAAATGTT.P142890-12	http://www.bolds...	14.10.2024	MachnoOceana	WMB0369-14	BOLD-AAA2226	96.99%	313	Arthropoda	Insecta	Lepidoptera	Hemiptera	Lepidoptera	Lepidoptera
AAATAAATGTT.P142890-13	http://www.bolds...	14.10.2024	MachnoOceana	WMB0369-14	BOLD-AAA2226	96.99%	313	Arthropoda	Insecta	Lepidoptera	Hemiptera	Lepidoptera	Lepidoptera
AAATAAATGTT.P142890-14	http://www.bolds...	14.10.2024	MachnoOceana	BEED089-15	BOLD-AAA2326	95.6%	313	Arthropoda	Insecta	Lepidoptera	Hemiptera	Lepidoptera	Lepidoptera
AAATAAATGTT.P142890-15	http://www.bolds...	14.10.2024	MachnoOceana	BEED089-15	BOLD-AAA2326	95.6%	313	Arthropoda	Insecta	Lepidoptera	Hemiptera	Lepidoptera	Lepidoptera
AAATAAATGTT.P142890-16	http://www.bolds...	14.10.2024	MachnoOceana	POLLE1841-19	BOLD-AAA2326	89.9%	303	Arthropoda	Insecta	Lepidoptera	Hemiptera	Lepidoptera	Lepidoptera

Sammlungsdatenverwaltung



Acc. Nr. **MAM2021-0661** **Crocodyra longipes Hutterer and Happold, 1983** Specimen 224844601 Event 1452635 Version 3 2 Withhold

2020/1/14 Gbama Community Forest
Altitude (mNN): 67 m NN
596.84 m W of Baima, Southern Province, Sierra Leone
Coordinates WGS84 (Longitude / Latitude) Long. (EW): -11.853072166 Lat. (NS): 8.109063148

MAM2021-0661
Crocodyra longipes Hutterer and Happold, 1983 [male]
Crocodyra longipes Hutterer and Happold, 1983 [respons: Hutterer, R.]
body length (incl tail) 1: 166 mm
head and body 1: 97 mm
tail 1: 69 mm
hind foot 1: 17 mm
ear 1: 12 mm
body weight 1: 18 g
Tissue Sample: ZFMK-TIS-59823
Bakar. I (No.: G120)
REL-224844601 (Duplicate)

2020/1/14 Gbama Community Forest
Altitude (mNN): 67 m NN
596.84 m W of Baima, Southern Province, Sierra Leone
Coordinates WGS84 (Longitude / Latitude) Long. (EW): -11.853072166 Lat. (NS): 8.109063148
BBC-MAM2021-0661
Crocodyra longipes Hutterer and Happold, 1983 [male]
Bakar. I (No.: G120)
Field-Number (Specimen-ID): G120

Collection specimen
Number: MAM2021-0661 Find next No.
Depositor: Dep. No.:
Source: Est. ID:
Acc. date: 9_7_2021 Suppl.: Cat.:
State: Withhold reason:

Exsiccata series:

Projects: Section Mammalia
Notes: Original: Sherman
Additional:

Create date: 2022-03-29 by Decher, Jan, Dr.
Last changes: 2022-04-25 by Collection Accession Unit, ZFMK

MAM2021-0661
MAM2021-0661-af - H12
Crocodyra longipes Hutterer and Happold, 1983 [male]
MAM2021-0661-ak - H323
Crocodyra longipes Hutterer and Happold, 1983 [male]

Sammlungsdatenverwaltung



The screenshot displays a specimen management interface with several panels:

- Top Panel:** Shows specimen details for *Crocidura longipes* Hutterer and Happold, 1983. It includes fields for Specimen Number (224844601), Event (1452635), Version (3), and Withhold status.
- Left Panel:** A tree view showing the specimen's location: 2020/1/14 Gbama Community Forest, Altitude (mNN): 57 m NN, 596.84 m W of Baine, Southern Province, Sierra Leone. It lists various morphological measurements like body length, head and body length, tail, hind foot, ear, and body weight.
- Center Panel:** A list of DNA metabarcoding results for *Enoplognatha*. It shows species names such as *Enoplognatha (jepons. consensus taxa)*, *Enoplognatha (jepons. consensus BOLD+NCBI+RDP)*, and *Enoplognatha (jepons. consensus NCB+RDP)*. It also includes a list of other species identified in the sample, such as *Entomobrya multifasciata*, *Calliphora vomitoria*, and *Archips infestans*.
- Right Panel:** A collection event summary for BPT_forest1, dated 2021-06-01 to 2021-08-31. It notes a 15-day sampling interval and provides details about the location (near the village of Repobolochen) and the collection method (blue pantraps).
- Bottom Panel:** A list of DNA metabarcoding results for *BPT_forest1*, showing various species like *Enoplognatha (jepons. consensus taxa)*, *Enoplognatha (jepons. consensus BOLD+NCBI+RDP)*, and *Enoplognatha (jepons. consensus NCB+RDP)*.

Sammlungsdatenverwaltung



The screenshots illustrate the software's capabilities in managing collection data, including:

- Specimen Entry:** Forms for recording specimen numbers (e.g., MAM2021-0661), collection dates, and locations.
- Geographic Data:** Integration with maps and coordinates (WGS84).
- Genetic Data:** Fields for DNA accession numbers and associated metadata.
- Taxonomic Classification:** Hierarchical selection of taxonomic groups (Kingdom, Phylum, Class, Order, Family, Genus, Species).
- Specimen Images:** A central image of a *Goniobranchus geometricus* specimen.
- Event and Collection Details:** Information about the collection event, including dates and sampling methods.

Datenverschneidung mit Laborsoftware



Import der Labordaten
aus Geneious als Analyse
(z.B. beim GBOL-Projekt)

Acc. Nr. ZFMK-TIS-11551
Specimen images and resources
Carabus nitens L., 1758
Specimen 11551 Event 55952 Version 22 5 Withhold. [locked]

Barcode_Analyse 161 (2015-03-02 10:54:50) 27120. 3 meth. (Barcode, ...)
Nr. of an.: 27120 Date: 2015-03-02 10:54:50
URI: [empty]
Part: [empty]
Respons.: GBOL-ZFMKLab
Result: AACCTTTATATTTTATTTTGGTGGCTGATCAGGAATAGTAGGAAC TTCATTAAGAATACTAATTCGAGCGGAATTAGGTAATCC TGGATCCTTAATGGAG ATGATCAAATTTATAATGTTATGTAACAGCCATGCTTTTGTAAATAATTTTTTATAGTTATGCCTATTATAATGGAGGATTTGGAAATGGATTAGTT CCTTTAATATTAGGAGCTCCGTGATAGCATTCCCTCGAATGAATAACATGAGATTTTGAACATTACCCCATCCTTAACCCCTCTCTTAATAGCAGTATA GTCGAAAAAGGGGCTGGGACAGGATGAACGTGTACCCCTCCGCTGCTTCGGGAATTGTCACAGAGGAGCATCAGTAGACCTAGCAATTTTCAGATTAC ATTTAGCAGGGATCTCTCAATTCAGGACAGTAATTTTATACGACAATTTAATAACGACATAGGAATAACCTTTGATGCAATACCATTTATTC GTATGATCTGTAGGATTTACAGCATTATTACTTTTACTTCTTTACCTGATTTAGCCGGAGCTATTACTATGCTTTTAAACAGATCGAAATTTAAATCTTCT TTTTTCGATCCGACGAGGAGGTGATCCTTACTTTTATCAACATTTATTT

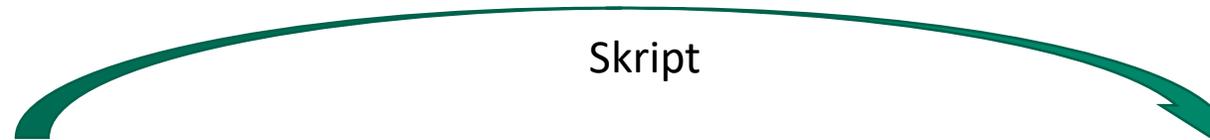
Notes: L2F insert

Methods
Barcode 9
Sequencing 2
Sequencing 3
project
region
sequence
sequence_length
trace_count
Parameter value
GBOL

Zusammenarbeit verschiedener Sektionen



Kopieren und Verknüpfung zwischen verschiedenen Sammlungen
(morphologische Sammlung ↔ Biobank)



Acc. Nr. MAM2021-0661
Crocidura longipes Hutterer and Happold, 1983

Collection specimen
Number: MAM2021-0661
Source: [empty]
Acc. date: 9. 7. 2021
State: [empty]

Specimen: 224844601
Event: 1452635
Version: 3 2
Withhold: [empty]

Projects
Section Mammalia

Notes
Original: Sherman

Sektion Mammalia

Acc. Nr. BBC-MAM2021-0661
Crocidura longipes Hutterer and Happold, 1983

Collection event. Number of specimen: 2
Date: 14. 1. 2020
End: [empty]
Ref.: 1
No.: [empty]
Country: Sierra Leone
Description of the locality: Gbama Community Forest
Methods

Specimen: 224847133
Event: 1452635
Version: 1 2
Withhold: [empty]

Projects
Section Mammalia

Notes
Original: Sherman

Sektion Biobank

Schnittstelle zu externen Datenbanken



Verlinkung zu extern gespeicherten Medien in fylr
(Voraussetzung: stabile und öffentliche URL)

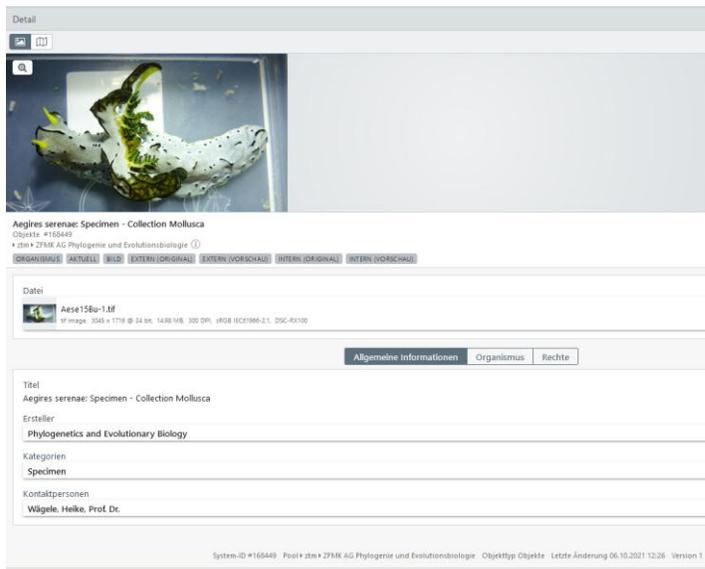
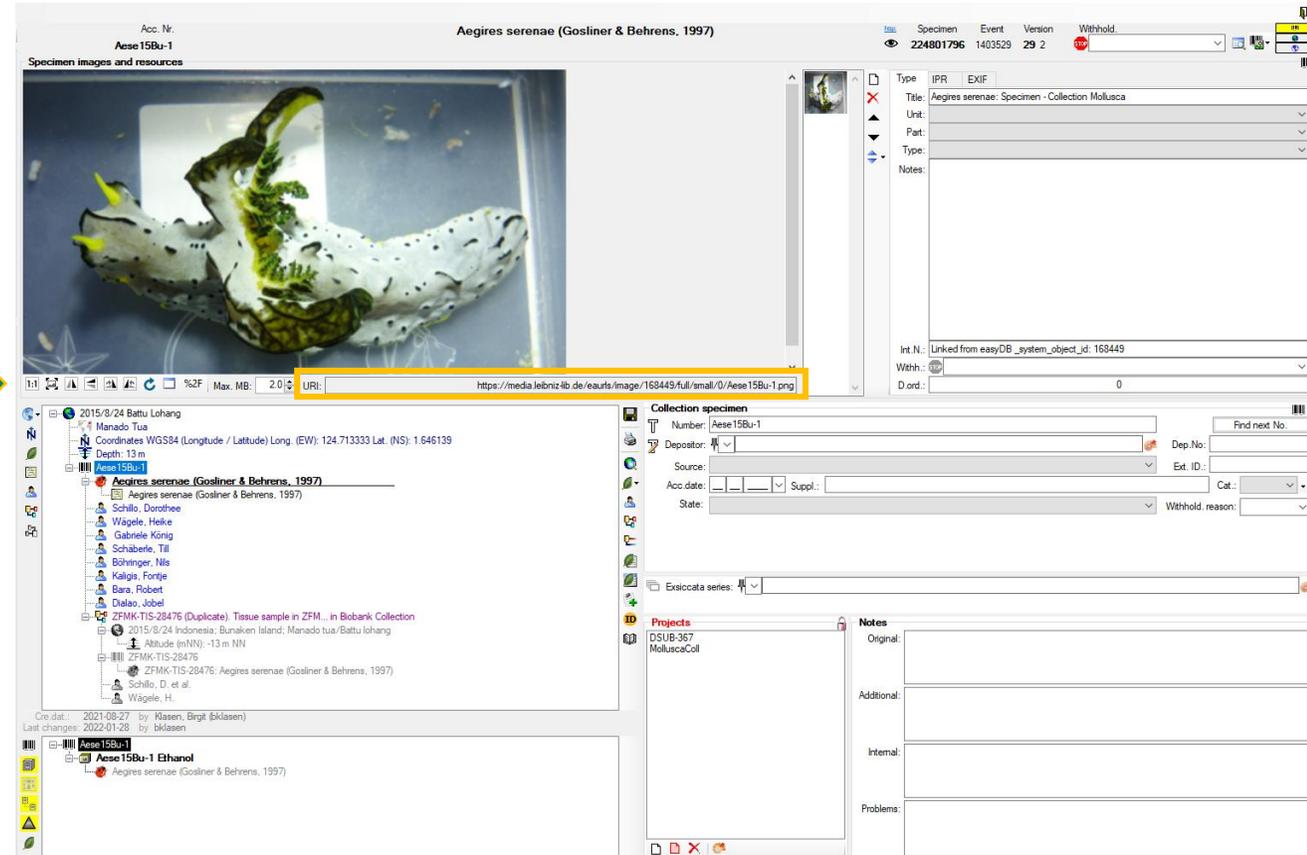


Bild in fylr



Link in DiversityCollection

Schnittstelle zu öffentlichen Repositorien



Verlinkung zu Sequenzen in ENA
(Voraussetzung: stabile und öffentliche URL)

ENA
European Nucleotide Archive

Project: PRJEB45552

General

View: XML (STUDY)

Download: XML (STUDY)

Read Files

Study Accession	Sample Accession	Experiment Accession	Run Accession	Tax Id	Scientific Name	Genera

Eintrag in ENA

Albinaria cretensis (Rossmässler, 1836)

Specimen images and resources

Collection specimen

Number: ZMH146002_4635

Source: ...

Acc date: ...

Identifier or URL of related specimen: <https://www.ebi.ac.uk/ena/browser/view/PRJEB45552>

External relation to Public database <https://www.ebi.ac.uk/ena/browser/view/PRJEB45552> in Malacology Section ZFMK

Link in DiversityCollection

Metadatenverwaltung in DiversityProjects



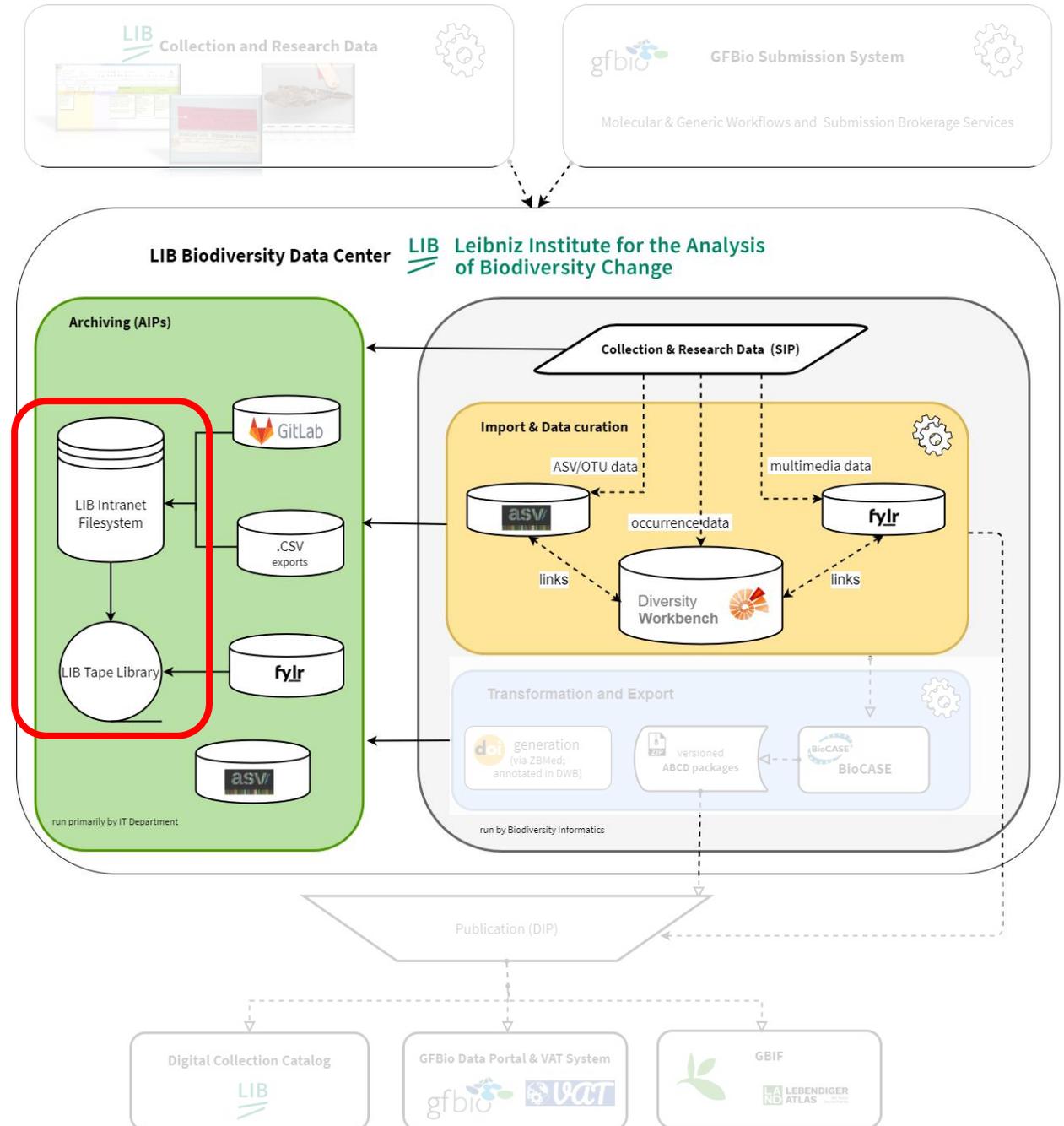
- Verwaltung projektweiter Metadaten
- Import der Metadaten bei GfBio-Projekten
- Einstellungen zur Veröffentlichung

The screenshot displays the DiversityProjects web interface for project DSUB-366. The interface is divided into several sections:

- Left Sidebar:** Contains navigation options such as 'Project', 'Description', 'URL', 'Notes', 'Settings', 'Agents', 'Licenses', 'Resources', 'References', and 'Descriptors'.
- Project Overview:** Shows the project name 'DSUB-366' and its ID '270066'. It includes a 'Title' field with the text 'First Study on Marine Heterobranchia (Gastropoda, Mollusca) in Bangka Archipelago, North Sulawesi, Indonesia'.
- Description:** A text area containing a detailed description of the study, mentioning the collection of 145 species in September 2017 and September 2018.
- Version and Embargo:** The version is set to '1' and the embargo date is '2021-10-31 00:00:00'. There are checkboxes for 'Create archive' and 'OAI P'.
- Contacts:** A detailed list of contacts is provided, including:
 - ContentContact:** Adelfia Papu, adelfia.papu@unsrat.ac.id
 - TechnicalContact:** LIB Biodiversity Datacenter, dataset@leibniz-lib.de, 049 (0) 228912242
 - Owner:** ZFMK & UNSRAT, h.roegele@leibniz-zfmk.de, adelfia.papu@unsrat.ac.id
- Bottom Section:** A list of related projects and their metadata, including project names, URLs, and descriptions.

Archivierung

Backups als .csv-Dateien im LIB Intranet Dateisystem und auf Bändern

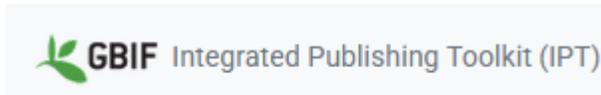


Publikation

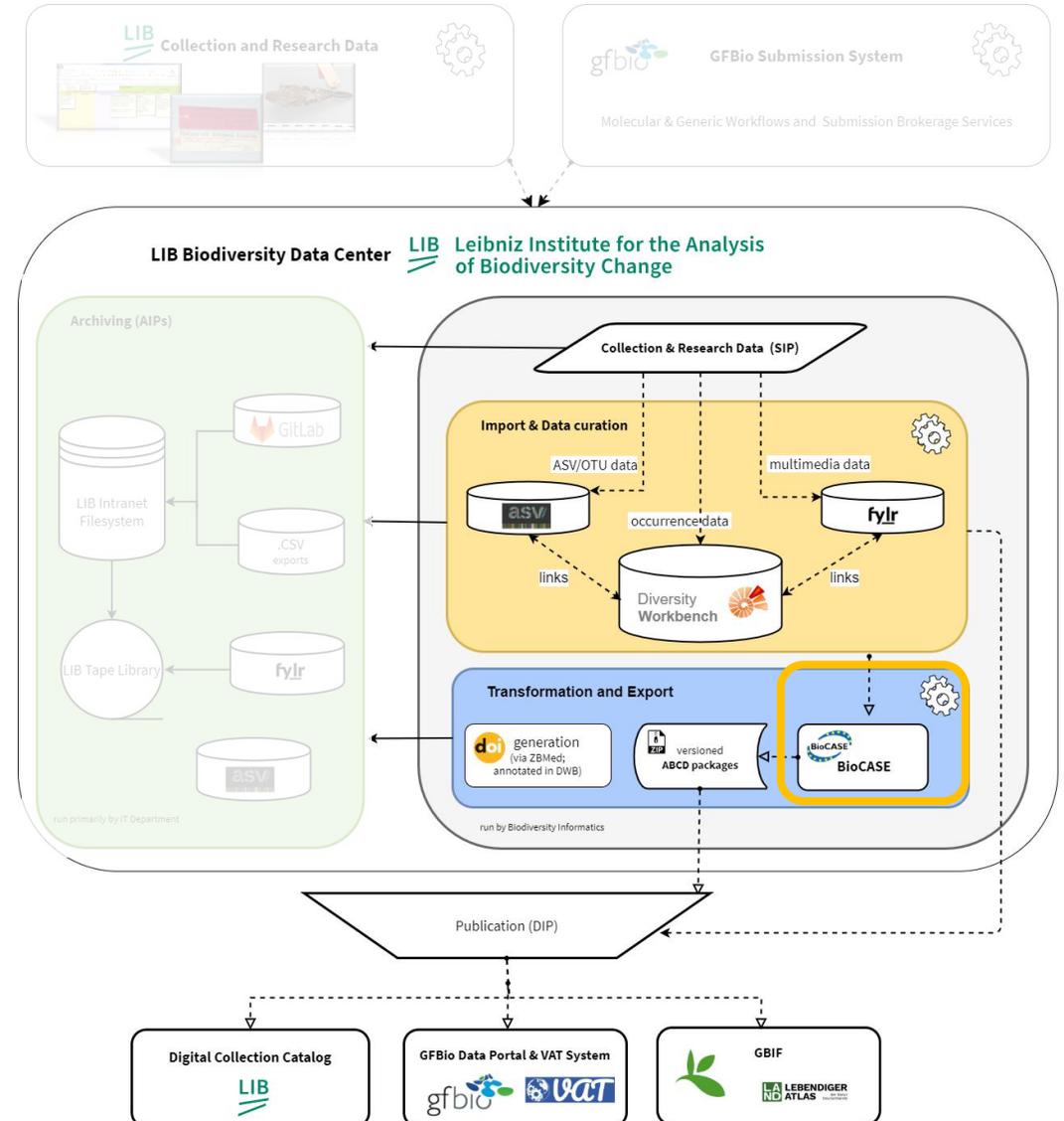
- Mapping zu ABCD in BioCASE Provider Software



- Views zum Mapping im GBIF IPT



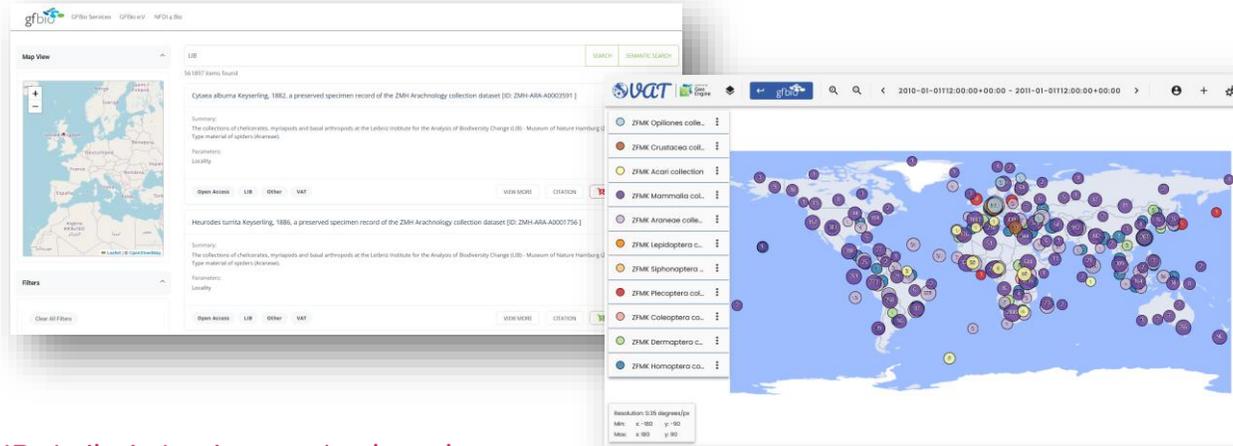
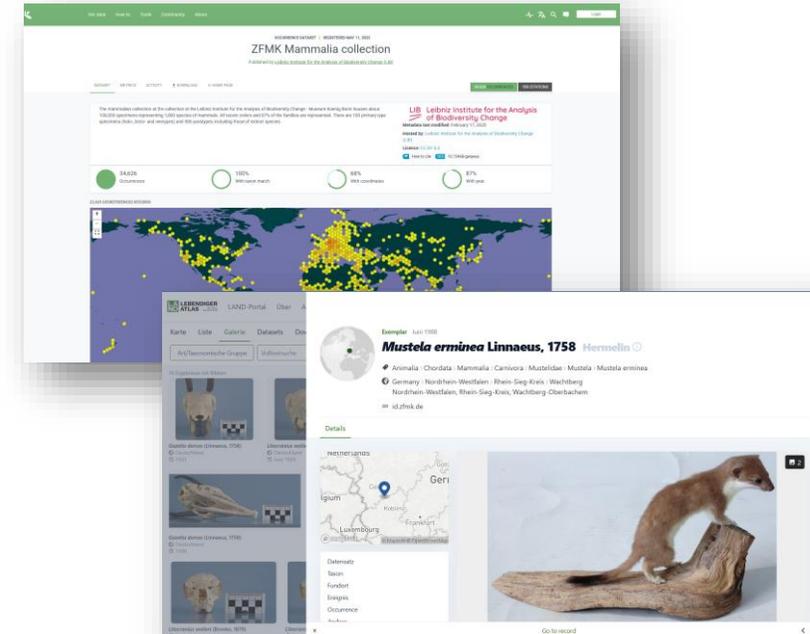
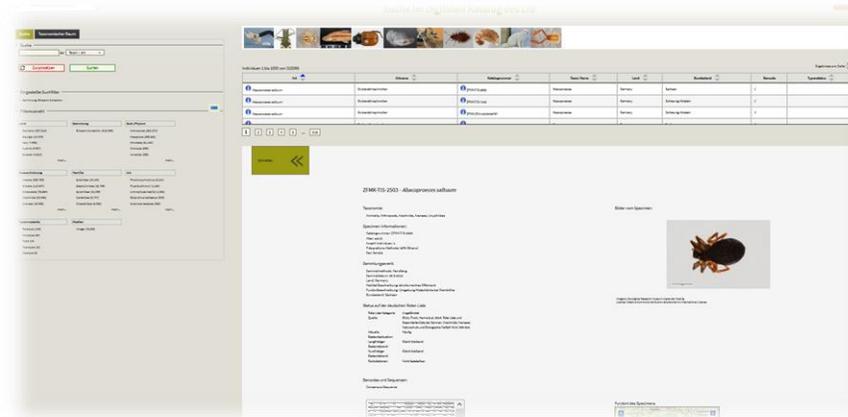
„Eine versteckte Sammlung ist eine tote Sammlung“



Recherche

„Eine versteckte Sammlung ist eine tote Sammlung“

- Daten werden sichtbar und durchsuchbar:
 - Digitaler Sammlungskatalog des LIB
 - GBOL-Portal
 - GfBIO-Portal/VAT
 - GBIF
 - GGBN



Vielen Dank für Ihre Aufmerksamkeit

Kontakt:

Name: Birgit Rach

E-Mail: b.rach@leibniz-lib.de