



BEXIS2 Forschungsdatenmanagement in Verbundprojekten

19.07.2023

Roman Gerlach

Friedrich-Schiller-University Jena (Germany)
Heinz Nixdorf Chair for Distributed Information Systems
Research Data Management Helpdesk

Intro

BEXIS2 is:

- designed for small to large **collaborative projects**
- focus on **active data** (i.e. project life-time)
- focus on **tabular data**, but not limited to
- focus on **data integration** and **re-use**
- generic, scalable, modular, **free** and open source

End-User Features



Search Data

Find datasets by keyword (incl. autocomplete) or use facets, filtering, sorting to quickly retrieve the dataset you need.



Upload Data

Upload your tabular data or any other type of data to BEXIS 2 and describe it with metadata.



Dataset Permissions

Specify fine grained data permissions on who can view, access, or update your dataset



Dataset Versioning

Keep track of the evolution of your dataset (i.e. who did what and when) and return to any previous version if needed.



Re-use Data Structures

Design your dataset by specifying its data structures, variables, or units only once in BEXIS 2 and re-use them for any dataset.



Single Sign-on

Use your existing credentials of your organisation to log on to BEXIS. No need to remember another username and password.



Web API Access

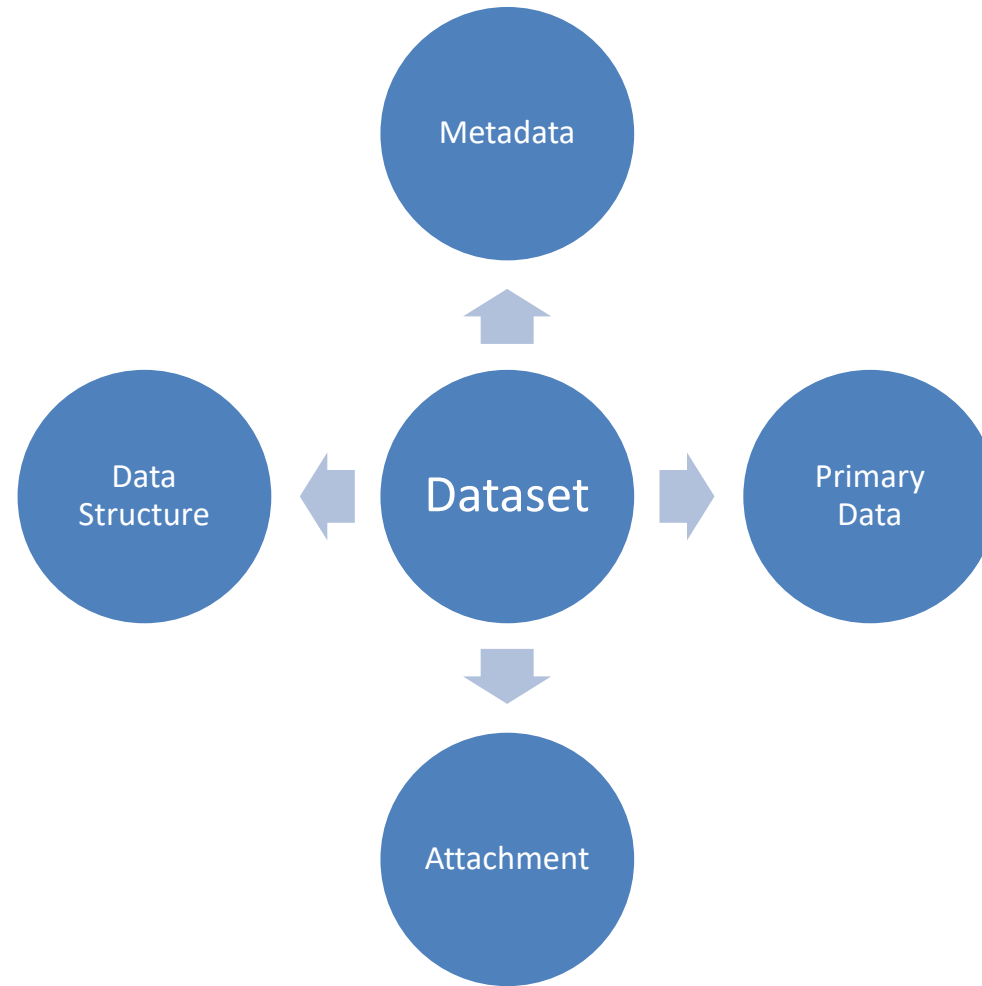
Beside downloading a dataset from the portal, you may also retrieve data stored in BEXIS directly with a script or from your application (e.g. R) using API calls.



Data Publication

Publish your dataset with a renown data repository or data journal (e.g. GFBio, Biodiversity Data Journal) and make it a citeable data publication identified by a DOI.

Concept: Dataset



Data in BEXIS 2

Tabular

- MS Excel, CSV, TSV

	A	B	C	D	E	F	G	H	I
1	Year	Site	Date	DOY	Lat	Long	Rev	Suction	Males
2	2005	ACRE	17. Jun	169	40,47	-86,99	-0,07	0	0
3	2005	ACRE	24. Jun	176	40,47	-86,99	0,05	1	0
4	2005	ACRE	01. Jul	183	40,47	-86,99	0,17	0	0
5	2005	ACRE	08. Jul	190	40,47	-86,99	0,28	0	0
6	2005	ACRE	15. Jul	197	40,47	-86,99	0,40	2	0
7	2005	ACRE	22. Jul	204	40,47	-86,99	0,52	11	0
8	2005	ACRE	29. Jul	211	40,47	-86,99	0,63	75	0
9	2005	ACRE	05. Aug	218	40,47	-86,99	0,75	214	0
10	2005	ACRE	12. Aug	225	40,47	-86,99	0,87	370	0
11	2005	ACRE	19. Aug	232	40,47	-86,99	0,98	82	0
12	2005	ACRE	26. Aug	239	40,47	-86,99	1,10	15	0
13	2005	ACRE	02. Sep	246	40,47	-86,99	1,22	4	0
14	2005	ALBERTL	08. Jul	190	43,76	-93,20	0,28	0	0
15	2005	ALBERTL	15. Jul	197	43,76	-93,20	0,40	97	0
16	2005	ALBERTL	22. Jul	204	43,76	-93,20	0,52	0	0

File based

- Multimedia data
- Geospatial data
- ...



Heterogeneity

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Year	Site	Date	DOY	Lat	Long	Rev	Suction	Males	Max_t	Mean_t	Min_t	GrowingDD50	precipmm
2	2005	ACRE	17. Jun	169	40,47	-86,99	-0,07	0	0	31	20,0	10	129	34,3
3	2005	ACRE	24. Jun	176	40,47	-86,99	0,05	1	0	35	23,0	10	169	93,9
									0	33	24,0	12	188	22,8
									0	32	22,0	13	162	20,2
									0	30	24,0	13	179	13,3

SEV_SmallMammalData_v.5.25.2010.xls								
	A	B	C	D	E	F	G	H
1	Date	Site	Plot	Species	Weight	Adult	Comments	
2	2/5/2010	Deep Well		1 DIPO	13.2	y		
3	2/4/2010	Deep Well		1 CLEGAP	11.6	j		
4	2/5/2010	Rio Salado		1 DIPO	14.2	y		
5	2/5/2010	Rio Salado		2 PERO	10.1	y		
6	3/15/2010	Deep Well		1 DIPO	15.2	y	plot burned	
7	3/15/2010	Deep Well		2 DIPO	21.7	y	pregnant	
8	3/15/2010							
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								
24								
25								
26								
27								
28								
29								
30								
31								
32								
33								
34								
35								
36								
37								
38								
39								
40								
41								
42								
43								
44								
45								
46								
47								
48								
49								
50								
51								
52								
53								
54								
55								
56								
57								
58								
59								
60								
61								
62								
63								
64								
65								
66								
67								
68								
69								
70								
71								
72								
73								
74								
75								
76								
77								
78								
79								
80								
81								
82								
83								
84								
85								
86								
87								
88								
89								
90								
91								
92								
93								
94								
95								
96								
97								
98								
99								
100								
101								
102								
103								
104								
105								
106								
107								
108								
109								
110								
111								
112								
113								
114								
115								
116								
117								
118								
119								
120								
121								
122								
123								
124								
125								
126								
127								
128								
129								
130								
131								
132								
133								
134								
135								
136								
137								
138								
139								
140								
141								
142								
143								
144								
145								
146								
147								
148								
149								
150								
151								
152								
153								
154								
155								
156								
157								
158								
159								
160								
161								
162								
163								
164								
165								
166								
167								
168								
169								
170								
171								
172								
173								
174								
175								
176								
177								
178								
179								
180								
181								
182								
183								
184								
185								
186								
187								
188								
189								
190								
191								
192								
193								
194								
195								
196								
197								
198								
199								
200								

Bacteria-clone sequenzen - Microsoft Excel										
	D	E	F	G	H	I	J	K	L	M
1	Accession Nr	Query	Identity	species (cult)	NC BLAST	taxonomy	phylum 1	phylum 2	phylum 3	consens
2	NR_040972.1	82	78	Levilina saccharolytica		Bacteria; Chloroflexi; Anaerolineae; Anaerolineales;	Chloroflexi	Chloroflexi	Chloroflexi	Chloroflexi
3	AF110272.1	87	78	Sarcina ventriculi		Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridia	Firmicutes	Firmicutes	Ignavibacteria	unclassified
4	NR_041306.1	99	79	Syntrophorhabdus aromaticivorans		Bacteria; Proteobacteria; Deltaproteobacteria; Syntroph	Deltaproteob	Deltaproteob	Bacteroidetes	unclassified
5	AY960779.1	97	79	bacterium Ellin		Bacteria; Verrucomicrobia; Verrucomicrobiae; Verrucomicro	Lentisphaerae	Verrucomicro	Verrucomicrobia	unclassified
6	JQ733906.1	86	79	Litorilina aerophila		Bacteria; Chloroflexi; Caldilineae; Caldilineales; Caldilli	Chloroflexi	Firmicutes	Chloroflexi	unclassified
7	AJ224540.1	98	79	unclassified		Bacteria; candidate division OP3.	unclassified	Actinobacter	unclassified	unclassified
8	KC251739.1	92	79	Gaiella sp.		Bacteria; Actinobacteria; Rubrobacteridae; Gaiellales; (Actinobacter	Deltaproteob	unclassified	unclassified
9	AJ431246.1	86	79	Dehalococcoides sp.		Bacteria; Chloroflexi; Dehalococcoidia; Dehalococcoida	Chloroflexi	Actinobacter	Cyanobacteria	unclassified
10	NR_074919.1	83	80	Desulfarculus baarsii		Bacteria; Proteobacteria; Deltaproteobacteria; Desulfa	Deltaproteob	Deltaproteob	Deltaproteobacteria	Deltaproteobacteria
11	AJ224540.1	100	80	unclassified		Bacteria; candidate division OP3.	unclassified	Deltaproteob	unclassified	unclassified
12	FR798928.1	90	80	Pleurocapsa cf. concharum		Bacteria; Cyanobacteria; Pleurocapsales; Pleurocapsa.	Cyanobacter	Firmicutes	Proteobacteria	unclassified
13	AF047571.1	63	80	candidate division OP11		Bacteria	unclassified	Betaproteob	Chloroflexi	unclassified
14	NR_102462.1	89	80	Cylindrospermum stagnale		Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Cyl	Cyanobacter	Cyanobacter	unclassified	unclassified
15	GU187029.1	86	81	Acidobacteria bacterium		Bacteria; Acidobacteria.	Acidobacteri	Deltaproteob	unclassified	unclassified
16	GU167989.1	87	81	Leptospirillum sp.		Bacteria; Nitrospirae; Nitrospirales; Nitrospiraceae;	Nitrospirae	Nitrospirae	unclassified	unclassified
17	KC852078.1	86	81	Leptospirillum ferriphilum		Bacteria; Nitrospirae; Nitrospirales; Nitrospiraceae;	Nitrospirae	Nitrospirae	unclassified	unclassified
18	FJ405900.1	58	81	Firmicutes bacterium		Bacteria; Firmicutes.	Firmicutes	Actinobacter	Chloroflexi	unclassified
19	FJ405900.1	57	81	Firmicutes bacterium		Bacteria; Firmicutes.	Firmicutes	Actinobacter	Chloroflexi	unclassified
20	FJ405900.1	85	81	Firmicutes bacterium		Bacteria; Firmicutes.	Firmicutes	Actinobacter	Bacteria	unclassified
21	HF952921.1	96	81	Thermoactinomyces bacterium						

Heterogeneity

For example: 18,200 different variables in 856 datasets

	A	B	C	D	E	J
1	datasetid	name	type	unit	description	attribute
2	2663	Datum	date		PlotName	Date
3	2900	date	date			Date
4	2349	DATUM	date	-	date of forest inventory	Date
5	2623	Zaehlstart	character(16)	Uhr	date of counting	Date
6	2901	floweringstart	integerNumber		the date when the plant starts to flower	DayOfTheYear
7	2900	diameter_stem	date	mm	the diameter of the stem	
8	2923	SamplingDate	date			Date
9	3020	date	date	none	dd.MM.yyyy	Date
10	3100	Datum_Jahr	date			Year
11	3923	Date	character(16)		date of radiotracking event	DateTime
12	4021	Datum	date	DD.MM.YYYY HH:MM:SS	Date and time of a measurement.	DateTime
13	4823	Datum	date	dd-MM-yyyy		Date
14	6096	StartTime	date		Sampling Start Time	Time

→ mapped into ~80 Data Variables

Dataset creation

Dashboard

Create

Collect

Search

Plan

David Schoene

Help

Home

Edit dataset:

Title

Gis

Metadata

Metadata

Edit

Last Modified

Visibility

not implemented

Edit

Status

not implemented

Edit

Data Description

> 60 min

datastructure test (2)

datastructure test

Id	Name	Unit	Datatype
1	plot	none	string
2	plant	none	string
3	date	none	date

File Upload

> 06 min

Drag "n" drop some files here, or click to select files max file : 1024 mb

.csv

Test Dataset DAVID.csv

Plants-Species-2022.csv

File Reader

Validation

Test Dataset DAVID.csv

Dataset

File

FileReader

Datastructure


Value

PrimaryKey

Can not convert to (2)


Plants-Species-2022.csv


*Sneak peek on version 3.0

 **BEXIS**
RESEARCH DATA MANAGEMENT

8

Data structure suggestion

 Dashboard Create Collect ▾ Search Plan ▾

Dauidschoene ▾ Help ▾ 

Home > Structure Suggestion ?


Delimiter
comma ▾

Decimal
point ▾

TextMarker
doubleQuotes ▾

Variable Unit Description Missing Values Data

☐ selection support



Missing Values
 +

Total: 3 Found: 3 Skipped: 0

plot plant date species total

pl1 plant a 2022-11-01 bug 12

pl2 plant b 2022-11-01 butterfly 121

Controls

Selection

left mouse button

Drag

left mouse button down and drag

Select Row

double click left mouse button

Deselect

right mouse button click

Data Structure Design

Units

Create new Unit +

Search rows...

Name	Description	Abbreviation	Dimension	Data Types	Measurement System
none	If no unit is used.	none	none	<ul style="list-style-type: none">stringdecimaldatedatetimebooleandoubleintegernon-negative longnon-negative integernumberbooltexttime	Unknown
mole per kilogram		mol/kg	amount per mass	double	Metric
katal per kilogram		kat/kg	amount per mass time	double	Unknown
mole per kilogram and second		mol/(kg*s)	amount per mass time	double	Metric
katal		kat	amount per time	double	Unknown
square meter		m^2	area	double	Metric
square meter per kilogram		m^2/kg	area per mass	double	Metric
arearatio		m^2/m^2	area ratio	double	Unknown
ampere		A	current	double	Metric
ampere-hour		Ah	current time	double	Metric

« < 10 > »

Units

meter

Name *
meter

Abbreviation *
abbreviation is required

Description *

Data Types *
double X

Dimension *
length

Name	System Type	Description
Double	double	A real number

Name	Dpecification	Description
length	L(1,0)M(0,0)T(0,0)I(0,0)?(0,0)N(0,0)J(0,0)	m, cm, km

Measurement System
Unknown **Metric** Imperial Nautical Natural

Search rows...

Name	Description	Abbreviation	Dimension	Data Types	Measurement System
per mile		pm	dimensionless	double	Imperial

Data access

- **Download**

- Including filter, sort, select of primary data
- different formats (e.g. CSV, Excel)
- bundles of primary data, metadata, data structure

- **Web API**

- e.g. direct access from **R** (rBEXIS package)

Web API Data Access

Sample REST API calls: Data

- <http://www.name.com/api/data/6>
 - /api/data/6?header=id,name
 - /api/data/6?filter=(Grade>50 AND Grade <90)
 - /api/data/6?header=id,name&filter=(Grade>50)

Sample REST API calls: Metadata

- <http://www.name.com/api/metadata/6>

Admin Features



Customizable Search Interface

Customize the Search Interface, e.g. select the metadata elements to be indexed and shown, specify the facets, and determine the representation of search results.



Extensible Architecture

BEXIS 2 follows a modular design and can be integrated with external applications and services or extended with custom modules



Use Multiple Metadata Schemas


Import multiple metadata schemas (XSD) relevant to your user community. For each schema a custom form will be generated by BEXIS.



System Feature Permission

Specify fine grained permissions on system features and tools for individual users or groups.

Modules & Extensions



Home Software ▾ Resources ▾ Community ▾ Knowledge Base ▾ About ▾

Modules Overview

This page contains modules which have been developed by the BEXIS2 community. Currently, they are provided "as is" without any warranty or liability. Please use at your own risk.

Most of the modules are constantly under development and are not (yet) provided as releases that match the official BEXIS2 releases. We hope to get there in the future. For the time being, if you are interested in using a module please contact the individual provider and to learn about the current status and how to use the module.

Name	Description	Contributor	Status	Contact
Module Template	This Module is a template for a BEXIS2 module to extend the functionality of the system.	BEXIS2	Public	BEXIS2 core team
BEXIS2-VAT System Interface (BEXIS2-VAT)	This BEXIS2 module extends the system with an interface to the Geo Engine	BEXIS2	Public (v1.0)	BEXIS2 core team
Data Summarization (DSM)		AquaDiva	Intern	Data Management Team
Event Management (EMM)	Simple to customize tool to manage participants registration for events (e.g. workshops, meetings, ...). Based on a fully flexible registration form. No user account needed for registration. Update your registration before the deadline.	Biodiversity Exploratories	Public	Data Management Team

Full list of modules: <https://bexis2.github.io/software/modules>

Former member Management (ALM)		Biodiversity Exploratories	BEXIS2 Core (SAM)	
Land-use Intensity (LUI)	A tool to calculate the index of land-use intensity (LUI) according to Blüthgen et al. (2012), based on information from the land owners on mowing, grazing and fertilization (Vogt et al. 2019) and your selections via a user-friendly interface.	Biodiversity Exploratories	Public	Data Management

FAIR Data Principles Conformance Assessment

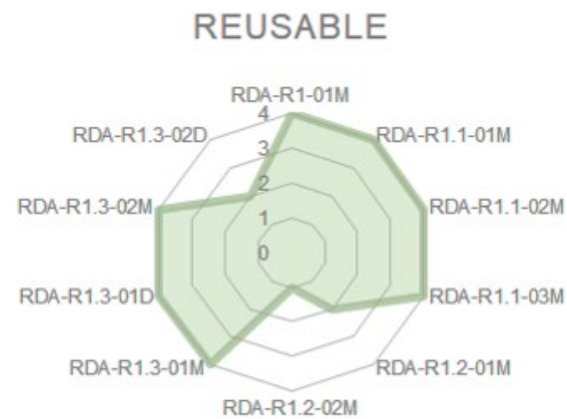
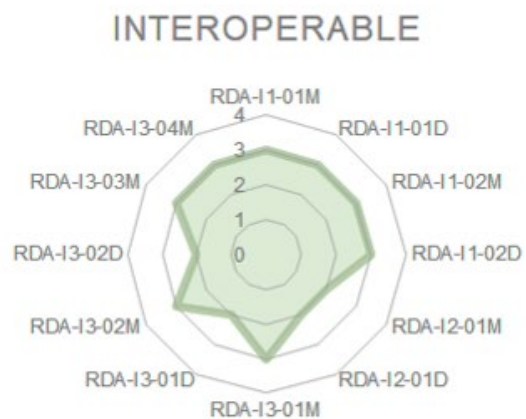
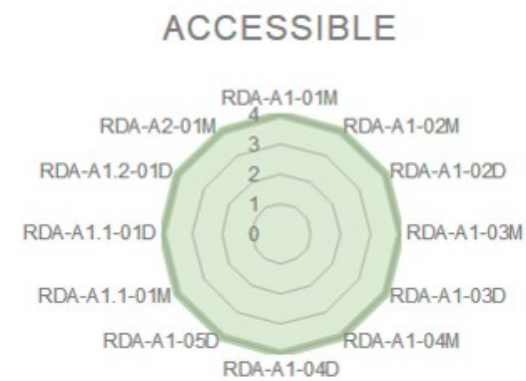
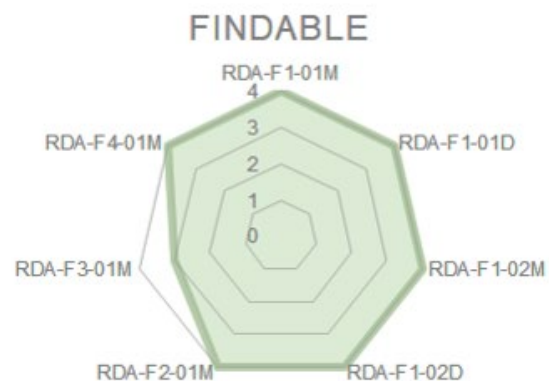


BEXIS2: A FAIR-aligned data management system for biodiversity, ecology and environmental data

▼ Javad Chamanara, Jitendra Gaikwad, Roman Gerlach, Alsayed Algergawy, Andreas Ostrowski, Birgitta König-Ries

<https://doi.org/10.3897/BDJ.9.e72901>

BEXIS2 FAIR conformance




Maturity level per indicator (per FAIR area) - 0 – not applicable; 1 – not being considered as yet; 2 – under consideration or in the planning phase; 3 – in implementation phase; 4 – fully implemented.

Typical BEXIS2 Deployments

Biodiversity
Exploratories
Information
System

DFG SPP 1374

<https://www.bexis.uni-jena.de/>


Biodiversity Exploratories Information System
Help

My Data Search Upload Tools Fieldwork Plots Documents Open Data Events

BEXIS - Search

* datasets only

Taxonomic Term

☐ Plants (896)
 ☐ None (601)
 ☐ Animals - Invertebrates (550)
 ☐ Microbes (398)
 ☐ Fungi (355)

more

Project

Consortia *

Grassland

☐ Yes (1255)
 ☐ No (1190)

Forest

Above Ground

Below Ground

☐ No (1521)
 ☐ Yes (924)

Process and Service *

Environmental Descriptor *

☐ Other (507)
 ☐ Soil (386)
 ☐ Land use (380)
 ☐ Habitat structure (293)
 ☐ Spatial (150)

more

Biotic Data Type *

Experimental Manipulation *

Data Type *

Type of Publication

Status of Publication

All

Search

1

2

3

4

5

6

7

8

9

10

20

Displaying items 1 - 20 of 2445

ID	Type	Title	Project	Start Date	End Date
31055	Publication	Interaction of plant roots and soil fungi and their effects on soil and rhizosphere enzyme activities	ECTOMYC	2019	2019
31054	Dataset	Cmic, Nmic, NH4+ and NO3- from SEG9 and SEG15 autumn 2017-summer 2018	BE_CH4	2017-11-09	2018-09-13
31051	Publication	Emission of CO2 and CH4 from 13 deadwood tree species is linked to tree species identity and management intensity in forest and grassland habitats	FunWood IV	2021	2021
31050	Publication	Sequential Monte-Carlo algorithms for Bayesian model calibration – A review and method comparison	CONNECT	2021	2021
31049	Dataset	Root dynamics in mineral topsoil across 150 forest plots in Exploratories (soil sampling campaign, 2017)	ECTOMYC	2021-07-21	2021-07-21
31048	Dataset	Root dynamics in organic layer across 150 forest plots in Exploratories (soil sampling campaign, 2017)	ECTOMYC	2017-05-01	2021-07-21
31047	Publication	Effects of global change on plants: tracing the footprints of climate warming and land use from herbaria to forest understories	PhenoDiv	2021	2021
31046	Publication	Land use intensification results in abrupt ecosystem changes in grasslands	Synthesis (core)	2021	2021
31045	Publication	Biodiversität im Wald	Knowledge Transfe...	2021	2021
31043	Publication	Microbial biomass and community structure predict rates of dark CO2 fixation in temperate forest soils	Soil abiotic (core)	2021	2021
31042	Publication	Short-term dynamics of structural complexity in differently managed and unmanaged European beech forests	Forest structure (c...	2021	2021
31041	Publication	Different community compositions between obligate and facultative oomycete plant parasites in a landscape-scale metabarcoding survey	PATHOGEN	2021	2021
31040	Dataset	Basic Information of all Forest Experimental Plots (FOX)	Instrumentation & r...	2020-10-06	2020-09-16
31039	Publication	Soil conditions drive belowground trait space in temperate agricultural grasslands	RootFun	2021	2021
31038	Publication	Global change and functional traits affect the importance of pest, intraguild and alternative prey for spiders in agricultural and non-agricultural habitats worldwide	SPIDERFUN	2021	2021
31037	Dataset	Coordinated Soil Sampling Protocol	Soil abiotic (core)	2021-05-03	2021-05-21
31036	Dataset	Greenhouse gas fluxes (CH4, CO2, N2O) from SEG9 and SEG15 autumn 2017-summer 2018	BE_CH4	2017-11-09	2018-09-13
31035	Dataset	60 Metatranscriptomes for two depths of SEG9 and SEG15 from November 2017 to October 2018	BE_CH4	2017-11-09	2018-10-13
31032	Dataset	Single sequence variants of cadaver-associated bacteria at forest EPs 2014	NecroPig	2014-08-04	2014-09-05
31031	Dataset	Single compound data of cadaver emitted volatiles at forest EPs 2014	NecroPig	2014-08-04	2014-09-05

1

2

3

4

5

6

7


8


9

10

20

Displaying items 1 - 20 of 2445

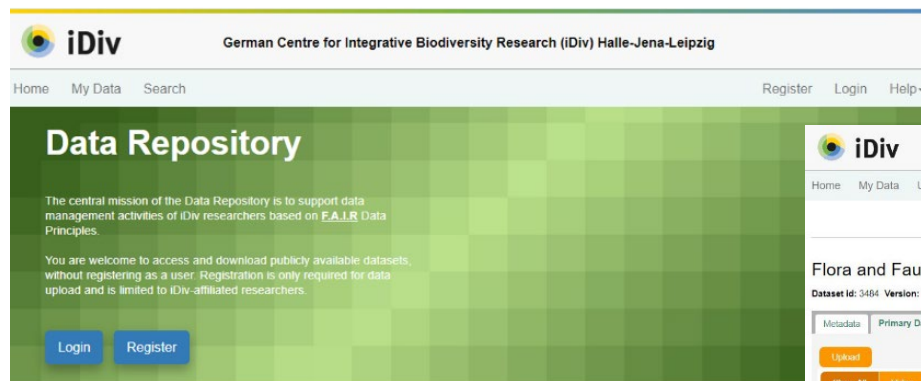

BEXIS
 RESEARCH DATA MANAGEMENT

Powered by

 Privacy Policy Terms and Conditions Contact Us Imprint

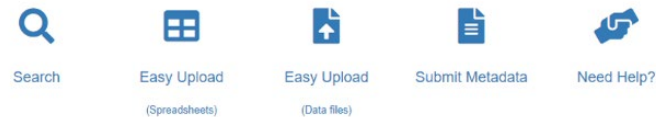
17

Typical BEXIS2 Deployments

Data Repository of the German Centre for Integrative Biodiversity Research (iDiv)
Halle-Jena-Leipzig



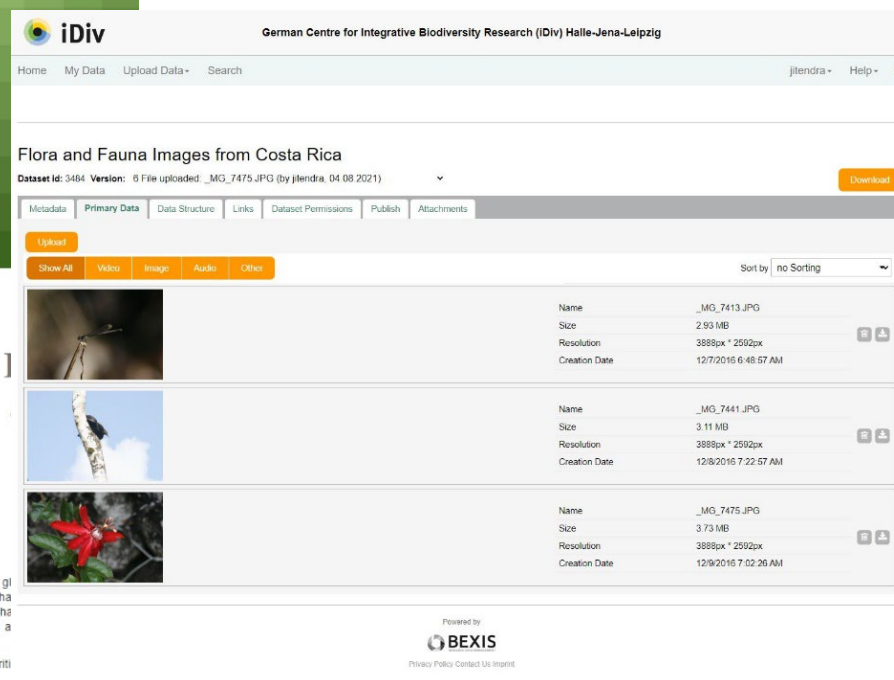
<https://idata.idiv.de/>



ABOUT

To promote theory-driven synthesis and data-driven theory, efficient management of data and information ranging from molecular to g1 required. In accordance with the [DFG research](#) data handling guidelines for biodiversity research, the Biodiversity Informatics Unit (BDU) has data repository for biodiversity and ecological data. The central role of the repository is to allow iDiv researchers to manage, describe, share, preserve biodiversity and ecology data, based on the Find, Accessible, Interoperable and Reusable (FAIR) data management principles a intellectual output of iDiv.

Some of the unique features of the repository are Easy Upload workflow for structured data. Generic Minimum Information Metadata writi



Community

Development Partner

FRIEDRICH-SCHILLER-
UNIVERSITÄT
JENA

Max Planck Institute
for Biogeochemistry



TUM
TECHNISCHE
UNIVERSITÄT
MÜNCHEN

GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN



PENSOFT®


GWGDG
Gesellschaft für wissenschaftliche
Datenverarbeitung mbH Göttingen

Project Instances



- Open Source Software Licence: GNU LGPL 3.0
- Code: <https://github.com/BEXIS2>

BEXIS2 Demo website

 Search

Register Login Help ▾

Demo ▾

Welcome to the BEXIS 2 demo.

The Bexis 2 demo can be explored in 3 ways.

Without Registration

As a non-registered user, you can view the publicly provided dataset.

Public Search

With prepared users

Role	User	Password
Scientist	scientist	scientist
Data Manager	datamanager	datamanager
Administrator	administrator	administrator

The system is reset once a week. Data that was created with this account no longer exists afterwards!

Login

Register yourself

You can register yourself. After successful activation, features rights must be granted by the administrator.

Please write an email to bexis-support@uni-jena.de and describe briefly what you want to test, so that we can activate the appropriate functions for you.

The system is reset once a week. Thus, the user accounts are no longer available and you have to register again!

Register

In the instance, there is test data that was provided to us by running instances.

Acknowledgement

THANKS to all who contributed towards the development of the BEXIS2 over the years...

Thanks to all developers:

Javad Chamanara, David Schöne, Sven Thiel, Martin Hohmuth, Nafiseh Navabpour, Thorsten Hindermann, Eleonora Petzold, Franziska Zander, Sirko Schindler, Masoud Allahyari, Payam Adineh, Arefeh Bahrami, Markus Steinberg, Michael Owonibi, Alsayed Algergawy, Hamdi Hamed



Thank you!

Questions?

Contact:

bexis-support@uni-jena.de

<https://bexis2.github.io/>