SENCKENBERG
world of biodiversity

edapho base

# Edaphobase – current status, plans and ambitions

Dr. Andrey S. Zaytsev

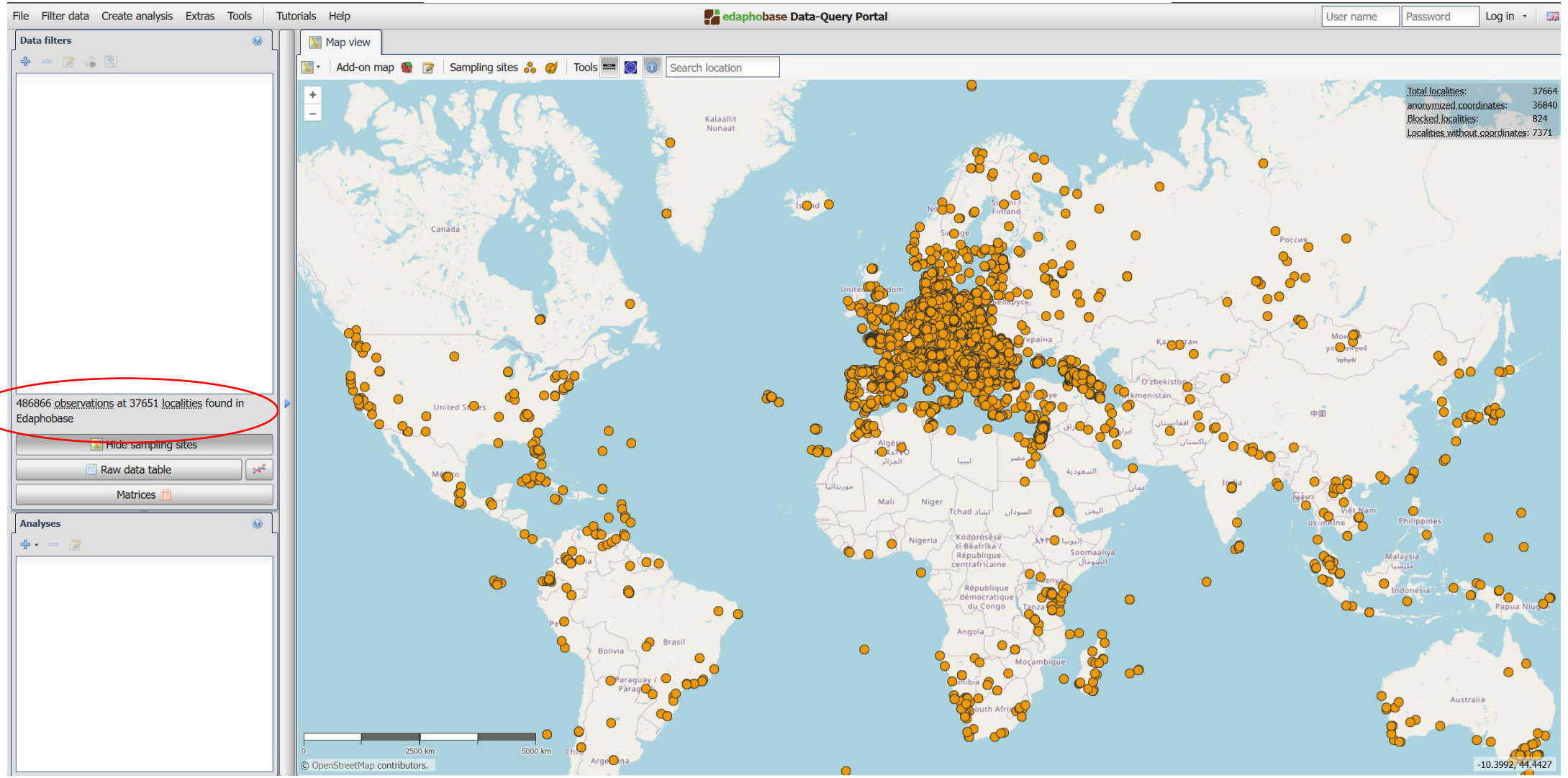
Section Soil Biodiversity Modelling,

Senckenberg Museum of Natural History Goerlitz

# Few facts

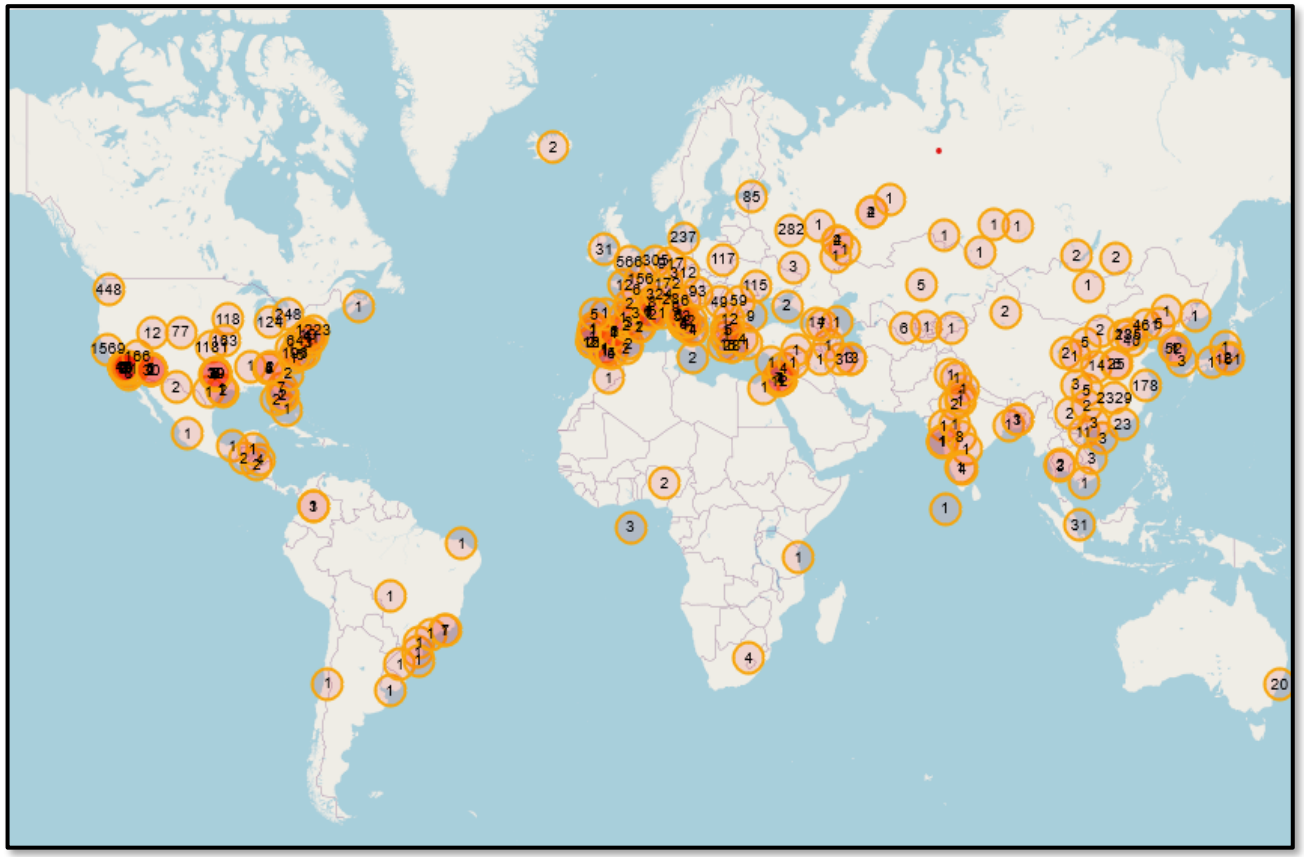
- Edaphobase is a non-commercial data infrastructure developed and hosted by the Senckenberg Museum of Natural History Görlitz in Germany.
- Edaphobase is a data provider to the Global Biodiversity Information Facility (GBIF).
- Initiated in 2009 by a group of enthusiasts inspired by Dr. David Russell.

# Edaphobase
## Current data distribution



# Edaphobase

## Global usage 2025 (log-ins to Edaphobase online portal)



**Removing spatial disbalance between data provision and acquisition:**

**GLOBAL AMBITION REQUIRING IMPROVED IMPORT PROCEDURE**

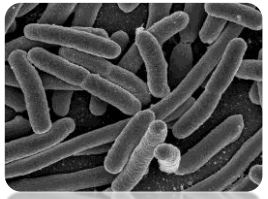
# Hosted Taxa

**Crassiclitellata** (earthworms)
**Enchytraeidae** (potworms)
**Nematoda** (threadworms)
**Oribatida** (moss/beetle mites)
**Gamasina** (gamasid mites)
**Collembola** (springtails)
**Chilopoda** (centipedes)
**Diplopoda** (millipedes)
**Isopoda** (woodlice)

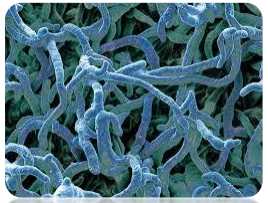
# Development I: More taxa welcomed

**Tardigrada**

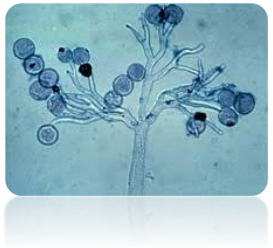
**Bacteria**



**Fungi**



**Oomycetes**



**+ functional parameters**

**(e.g. basal respiration)**

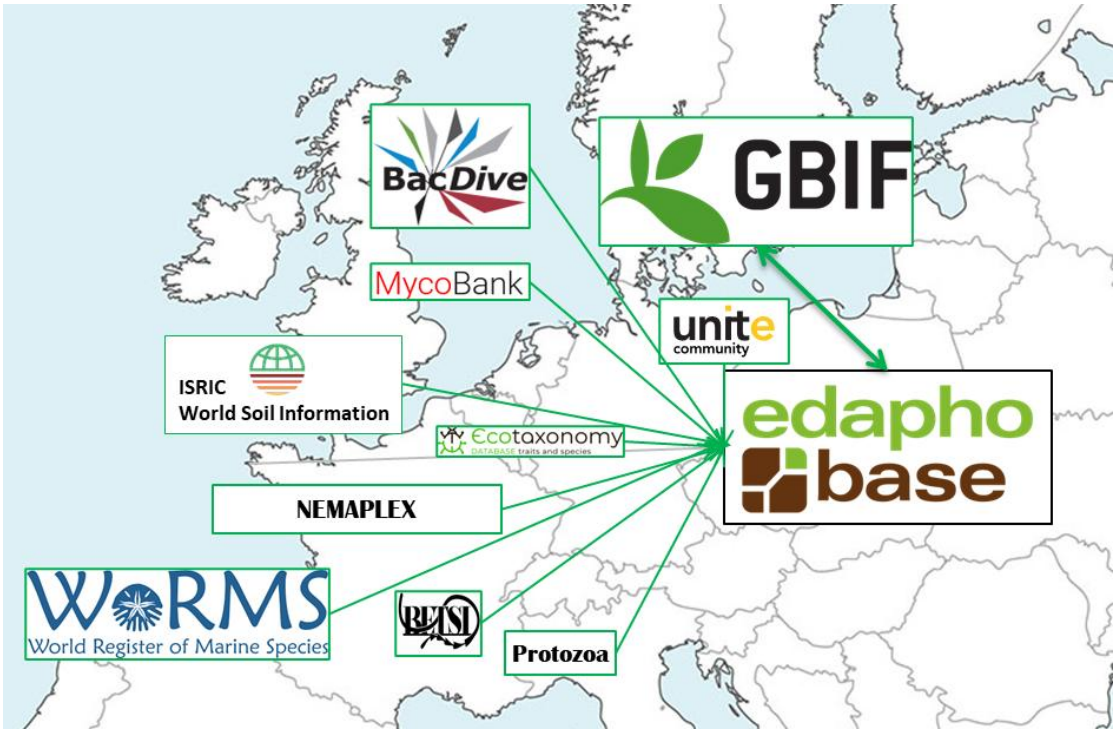
# Development II: more active in-house data digitalization

| Taxonomic group | Data sources | covered taxa | mentioned locations | data records |
|---|---|---|---|---|
| Lumbricidae (literature) | 135 | 95 | 807 | 9.389 |
| Lumbricidae (collection) | 648 | 74 | 451 | 651 |
| Lumbricidae (raw data) | 19 | 58 | 478 | 10.021 |
| Lumbricidae (total) | 802 | 150 | 1.709 | 20.061 |
| Enchytraeidae (literature) | 31 | 125 | 128 | 4.309 |
| Enchytraeidae (collection) | 1 | 1 | 1 | 1 |
| Enchytraeidae (raw data) | 7 | 126 | 87 | 8051 |
| Enchytraeidae (total) | 39 | 156 | 236 | 12.361 |
| Nematoda (literature) | 216 | 1.311 | 712 | 36.718 |
| Nematoda (collection) | 1.630 | 506 | 1.120 | 1.630 |
| Nematoda (raw data) | 1 | 145 | 55 | 1.982 |
| Nematoda (total) | 1.846 | 1.706 | 1.830 | 38.348 |
| Collembola (literature) | 246 | 1.225 | 1.537 | 80.809 |
| Collembola (collection) | 8.492 | 565 | 840 | 11.885 |
| Collembola (raw data) | 17 | 273 | 346 | 12.331 |
| Collembola (total) | 8.755 | 1406 | 2.691 | 105.025 |
| Oribatida (literature) | 16 | 417 | 198 | 7.629 |
| Oribatida (collection) | 23.718 | 1.392 | 1.323 | 25.382 |
| Oribatida (raw data) | 6 | 274 | 59 | 11.202 |
| Oribatida (total) | 23.740 | 1.519 | 1.578 | 44.213 |

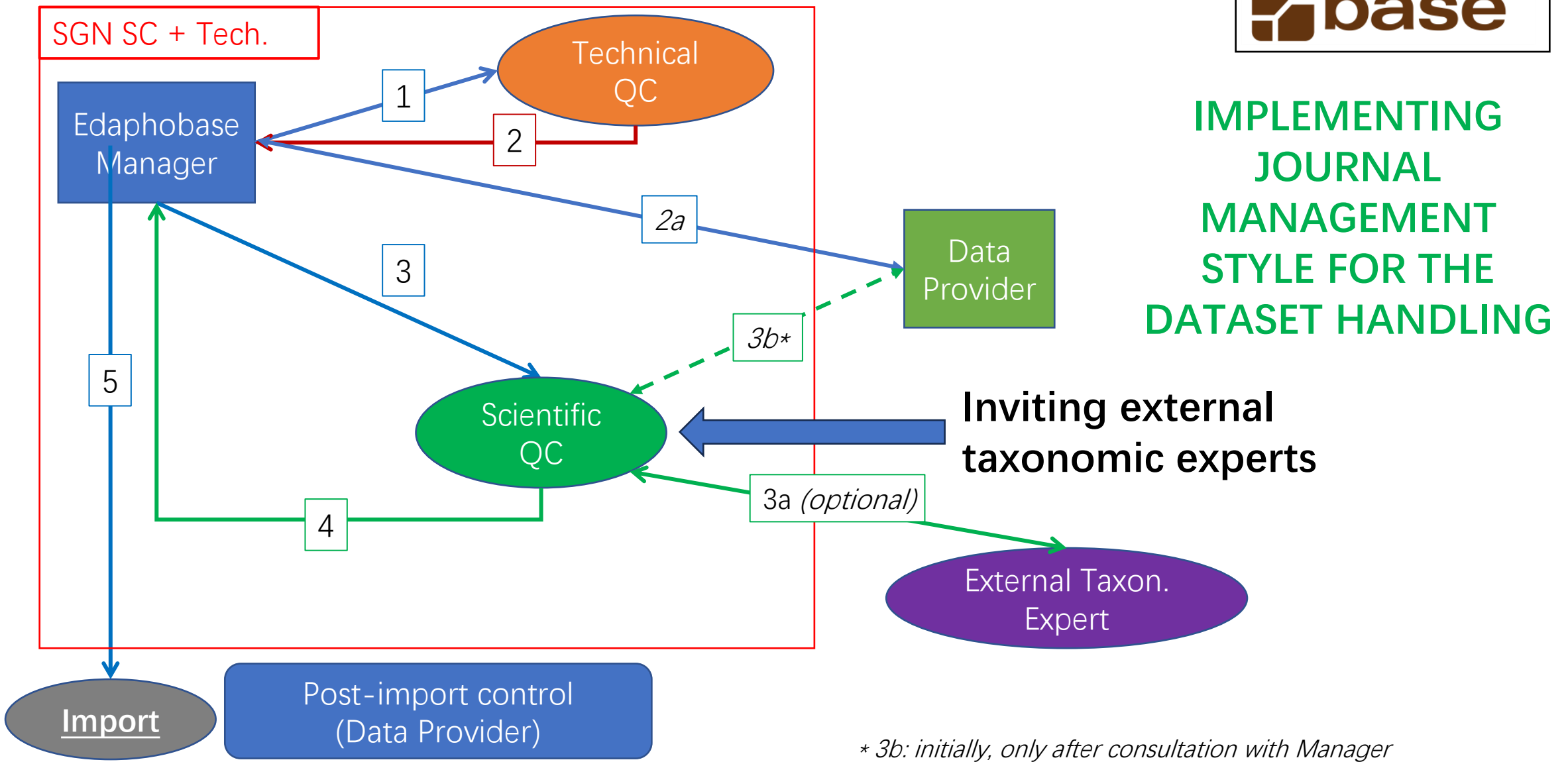
# Development II: more active in-house data digitalization

| Taxonomic group | Data sources | covered taxa | mentioned locations | data records |
|---|---|---|---|---|
| Gamasina (literature) | 91 | 972 | 548 | 10.137 |
| Gamasina (collection) | 10.195 | 888 | 2.101 | 10.195 |
| Gamasina (total) | 10.286 | 1.493 | 2.642 | 20.332 |
| Chilopoda (literature) | 1.403 | 1.611 | 1.650 | 14.890 |
| Chilopoda (collection) | 25.927 | 752 | 5.165 | 26.759 |
| Chilopoda (raw data) | 12 | 34 | 54 | 798 |
| Chilopoda (total) | 27.342 | 2008 | 6.827 | 42.447 |
| Diplopoda (literature) | 1.703 | 2.002 | 3.052 | 21.980 |
| Diplopoda (collection) | 48.374 | 2.479 | 8.383 | 50.186 |
| Diplopoda (raw data) | 13 | 60 | 125 | 1.848 |
| Diplopoda (total) | 50.090 | 3.744 | 11.331 | 74.014 |
| Isopoda (literature) | 273 | 84 | 1.309 | 5.772 |
| Isopoda (collection) | 5.219 | 6058 | 1.824 | 5.219 |
| Isopoda (raw data) | 4 | 18 | 28 | 550 |
| Isopoda (total) | 5.496 | 94 | 2.845 | 11.541 |
| Other (literature) | 127 | 242 | 263 | 985 |
| Other (collection) | 5.563 | 858 | 1.708 | 5.563 |
| Other (raw data) | 2 | 102 | 20 | 2544 |
| Other (total) | 5.692 | 1.154 | 1.976 | 9.092 |

# Development III: more systemic and automated outsourcing of taxonomic backbone and environmental data maintenance

# Development IV: enhancing data peer review & QC

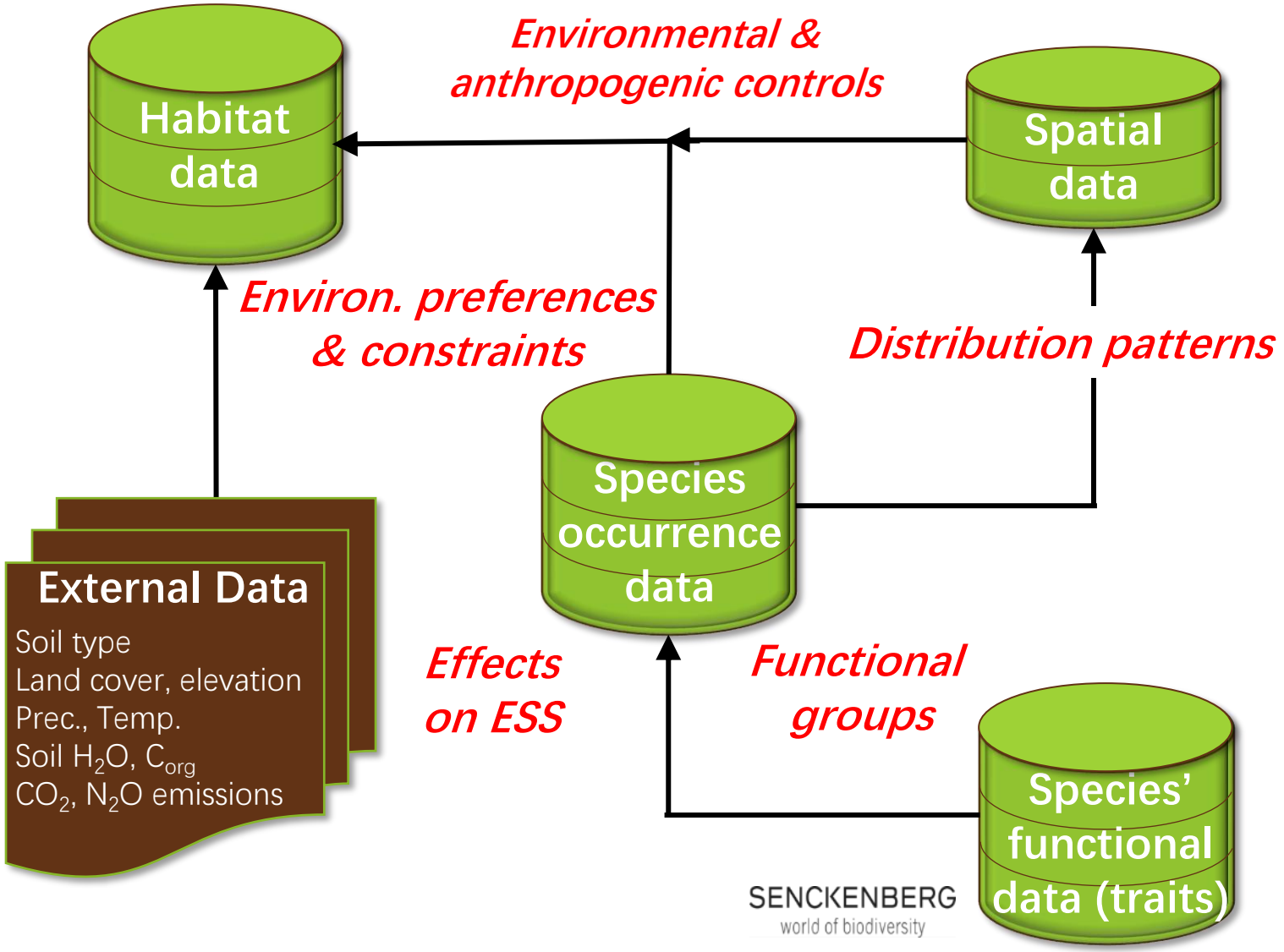



**IMPLEMENTING JOURNAL MANAGEMENT STYLE FOR THE DATASET HANDLING**

** 3b: initially, only after consultation with Manager*

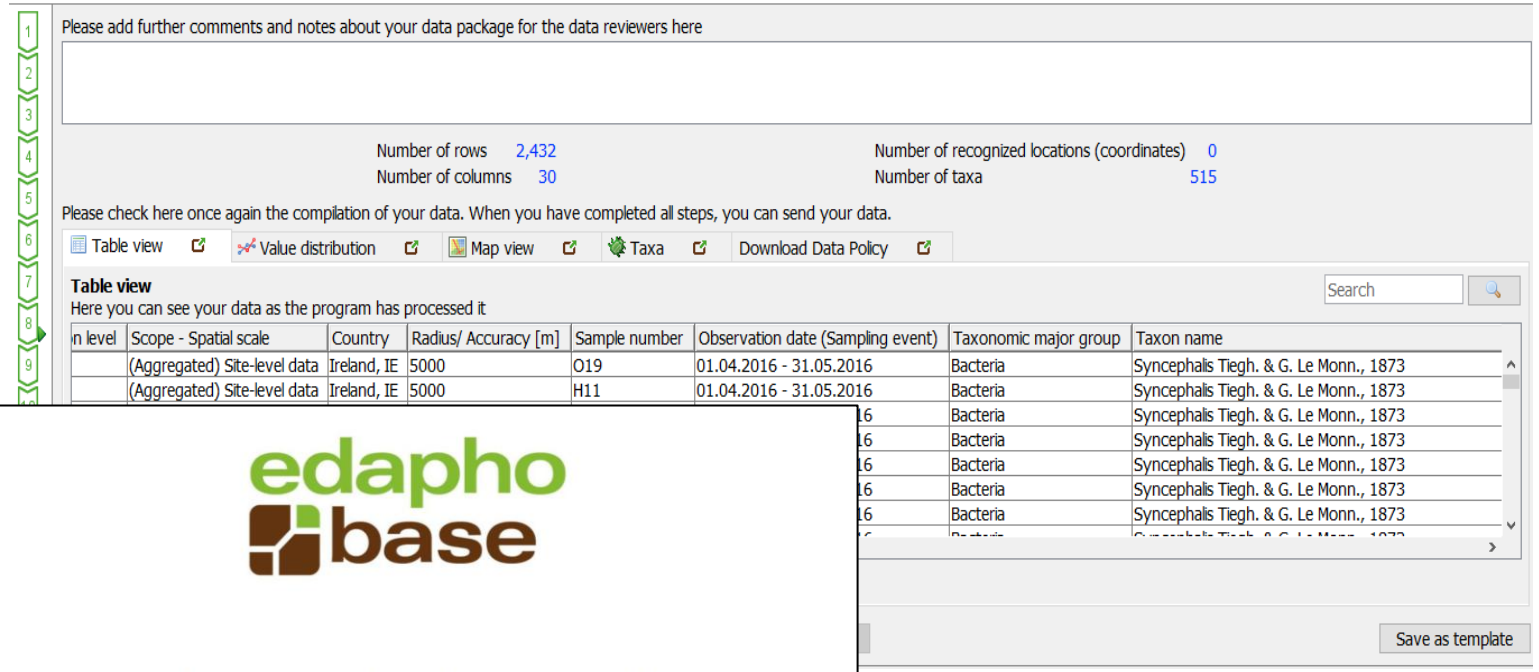
SENCKENBERG world of biodiversity

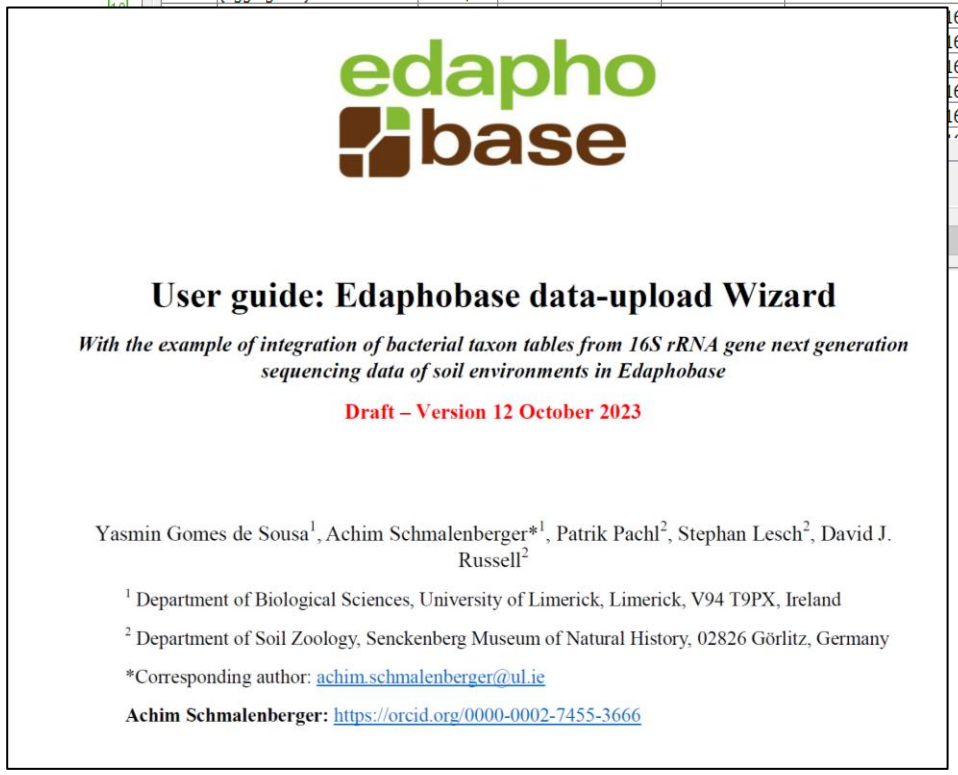
# Development V: From problems to solutions: data management and modelling




- Bridging taxonomic and functional data
- Linking soil fauna and environmental data
- Integrating advanced modelling (Edaphostat 2.0)
- Standardization and ontologies

# Development VI: Data Import Wizard



**User guide: Edaphobase data-upload Wizard**

*With the example of integration of bacterial taxon tables from 16S rRNA gene next generation sequencing data of soil environments in Edaphobase*

**Draft – Version 12 October 2023**

Yasmin Gomes de Sousa[1], Achim Schmalenberger*[1], Patrik Pachl[2], Stephan Lesch[2], David J. Russell[2]

[1] Department of Biological Sciences, University of Limerick, Limerick, V94 T9PX, Ireland

[2] Department of Soil Zoology, Senckenberg Museum of Natural History, 02826 Görlitz, Germany

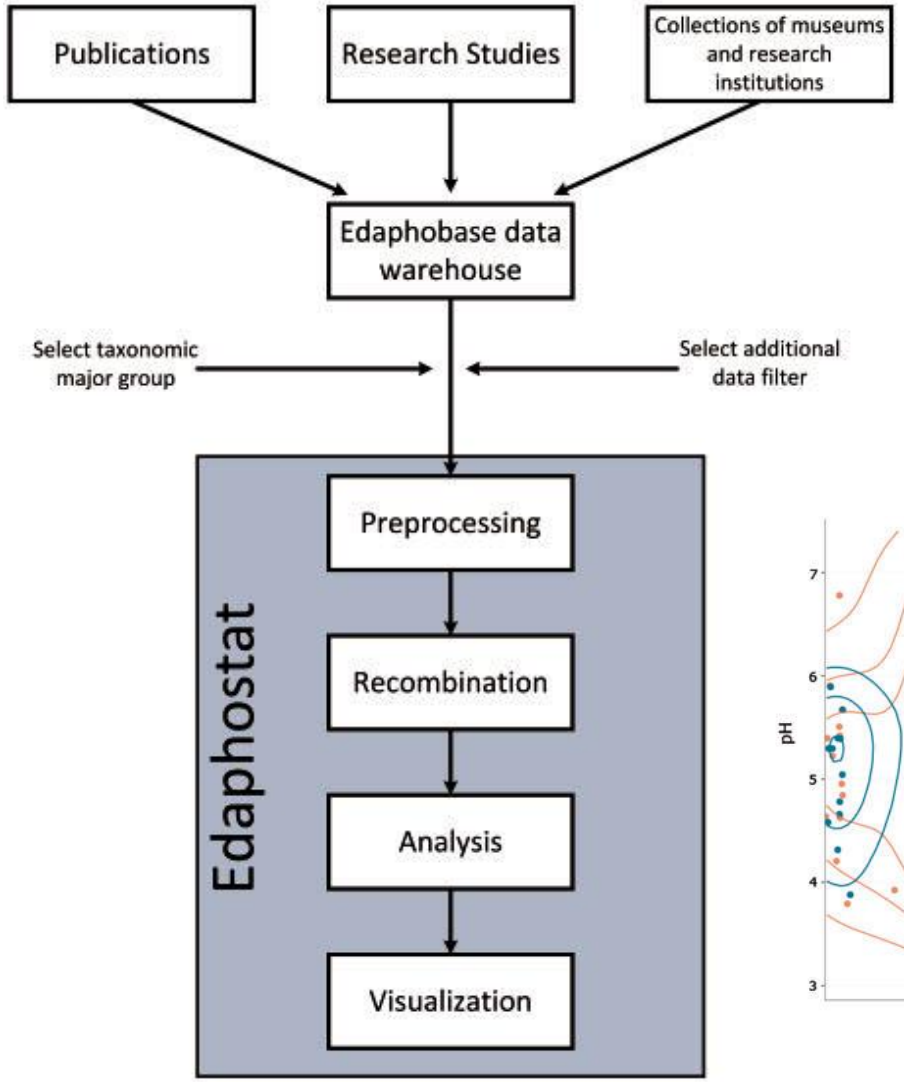
*Corresponding author: achim.schmalenberger@ul.ie

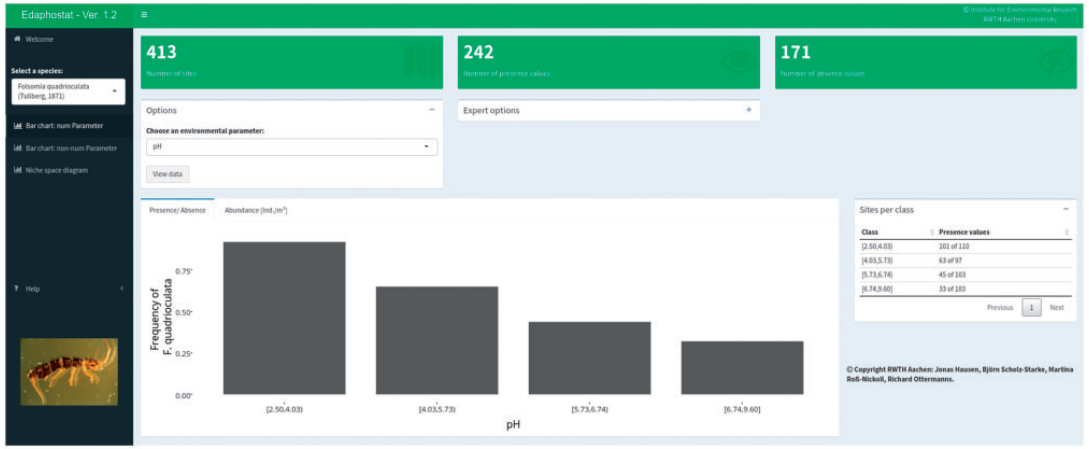
**Achim Schmalenberger:** https://orcid.org/0000-0002-7455-3666

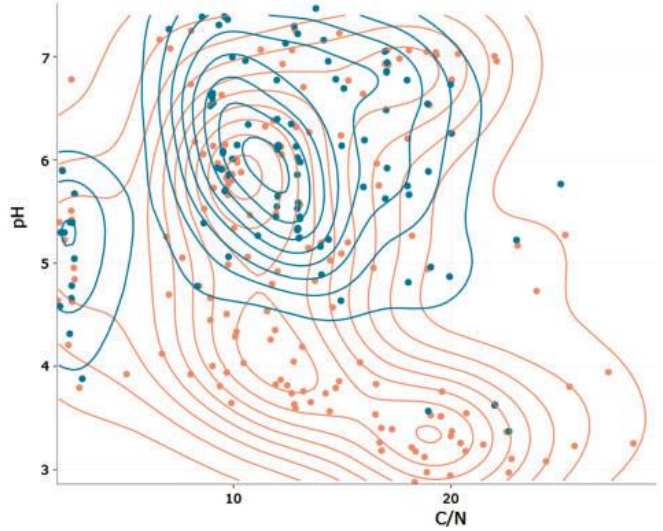
- Omnivorous in terms of data formats
- Many built-in data checking options
- BUT: Still hard to use without experience.

SENCKENBERG
world of biodiversity

# Development VII: In-house data analysis - Edaphostat




SENCKENBERG
world of biodiversity

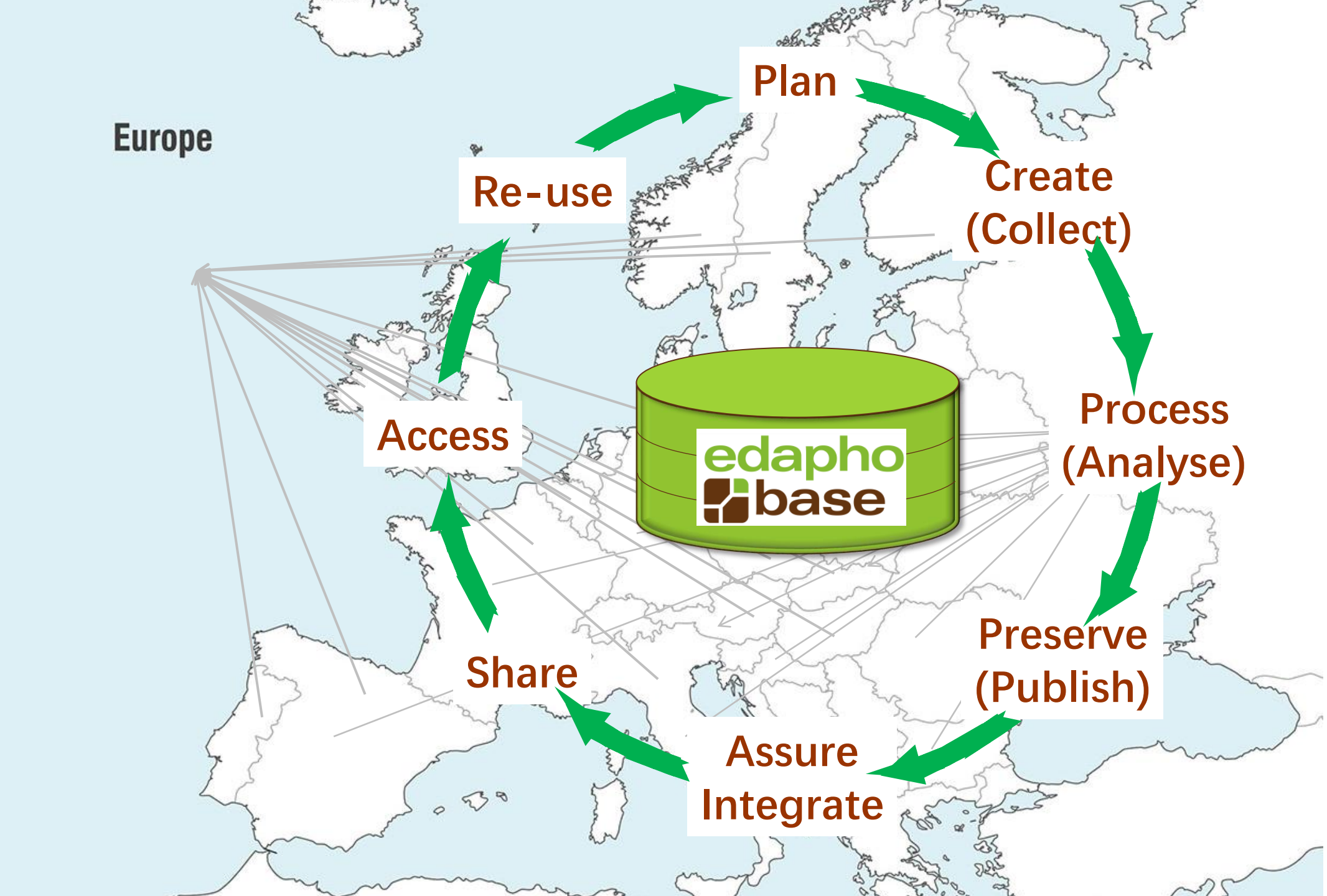
DATABASE
*Database*, 2017, 1–6
doi: 10.1093/database/bax080
Technical Report

Technical Report

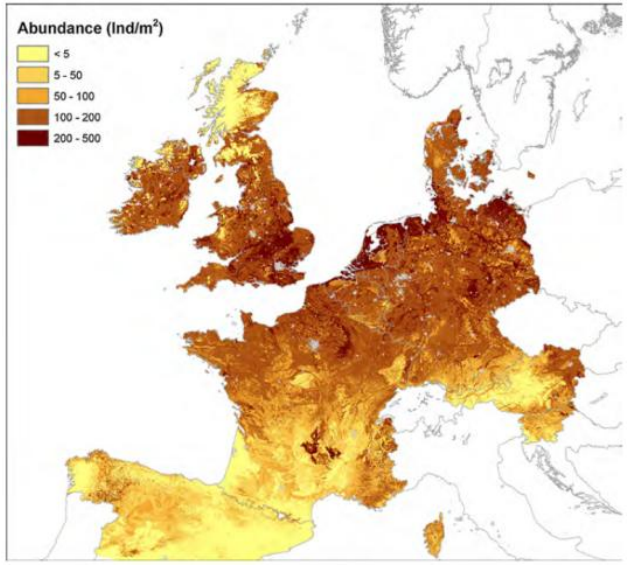
## Edaphostat: interactive ecological analysis of soil organism occurrences and preferences from the Edaphobase data warehouse

Jonas Hausen[1,*], Björn Scholz-Starke[1], Ulrich Burkhardt[2], Stephan Lesch[2], Sebastian Rick[2], David Russell[2], Martina Roß-Nickoll[1] and Richard Ottermanns[1]

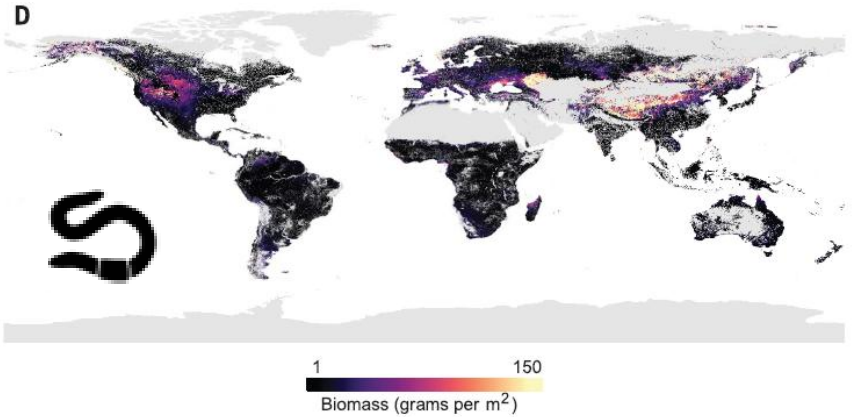
[1]Institute for Environmental Research, RWTH Aachen University, Worringerweg 1, 52074 Aachen, Germany and [2]Senckenberg Museum of Natural History Görlitz, P.O. Box 300154, 02806 Görlitz, Germany


Europe
Plan
Create
(Collect)
Process
(Analyse)
Preserve
(Publish)
Assure
Integrate
Share
Access
Re-use
edapho
base

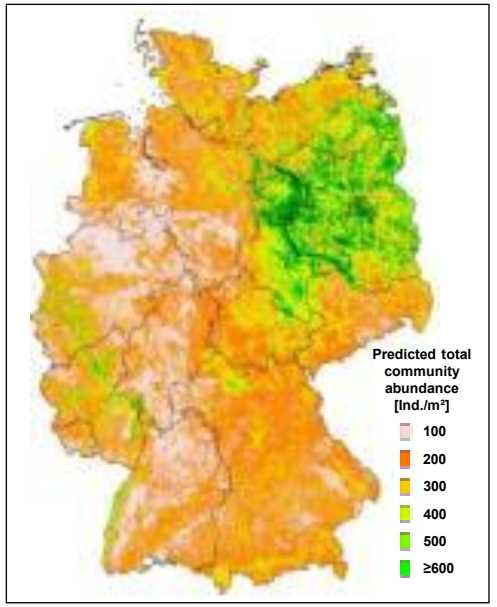
# Distribution modelling




Rutgers et al. 2016
*(Appl. Soil Ecol.)*

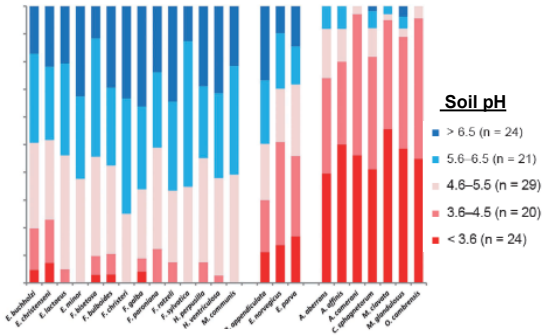



Phillips et al. 2019
*(Science)*




Salako et al. 2023
*(Biodiv. & Conserv.)*

# Niche space analyses




Römbke et al. 2013
*(Soil Org.)*

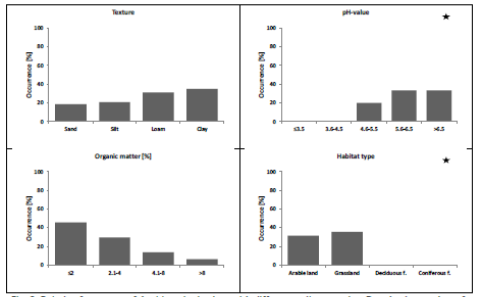
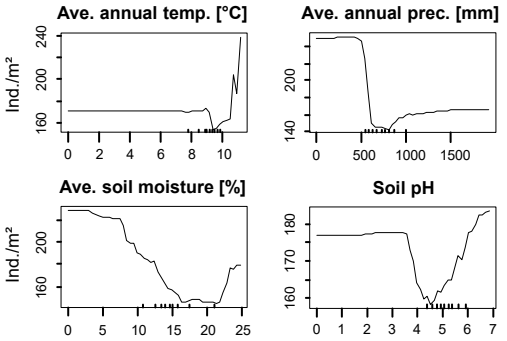
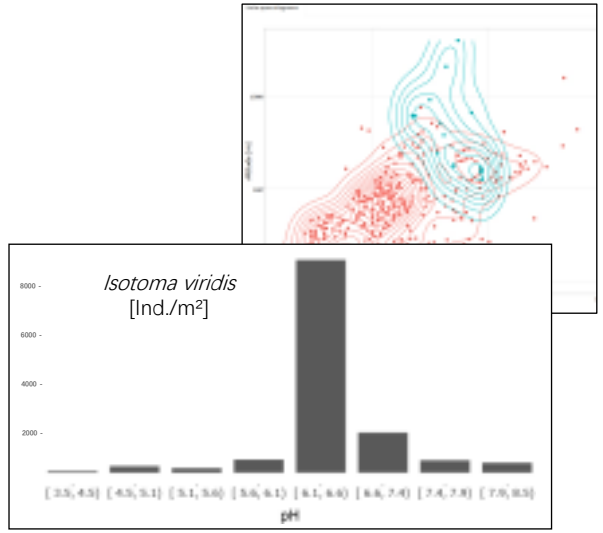



Fig. 2: Relative frequency of *A. chlorotica* in sites with different soil properties. Data basis: number of sites at which this species was found (Table 2). Star: statistically significant difference ($Chi^2$-Test).

Jänsch et al. 2013
*(Soil Org.)*



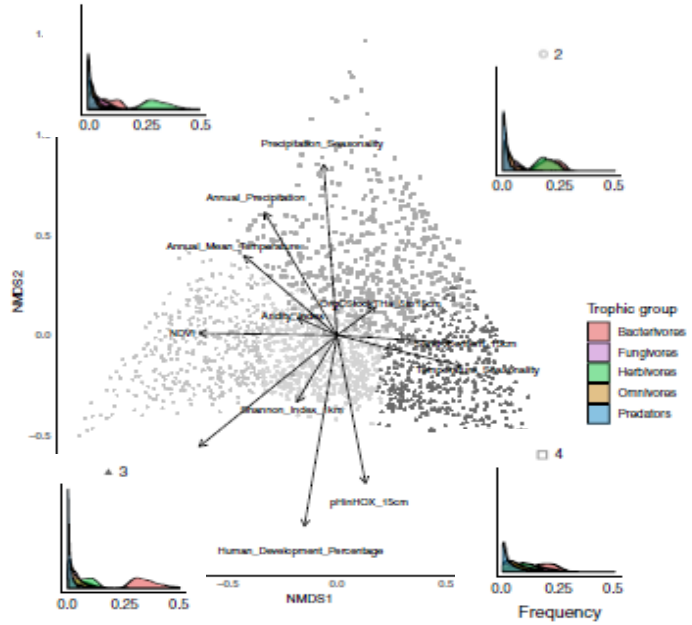

Salako et al. 2023
*(Biodiv. & Conserv.)*




Hausen et al. 2017
*(DataBase)*

# Soil Food Web Analyses

Marizidoviek et al. 2022
*(Ecol. Modell.)*

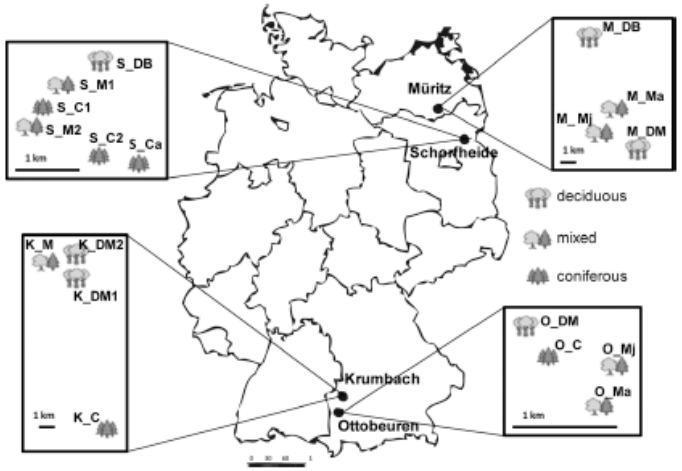



Hoogen et al. 2019
*(Nature)*

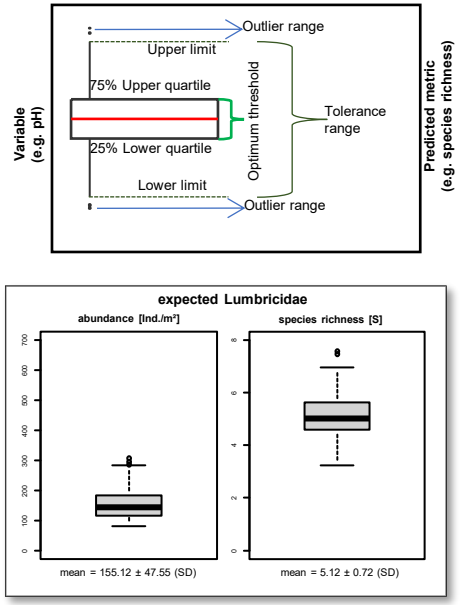
# Applied


Lehmitz et al. 2016

**Forest conversion sites**




Russell & Gergocs 2019
*(Forest Ecology & Management)*



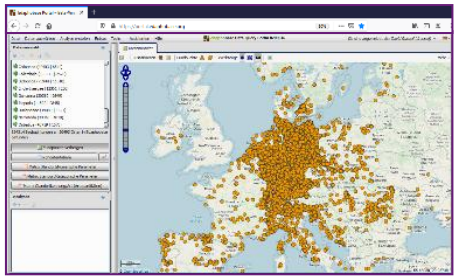

Russell & Salako *(in prep.)*

# Usecase I: Knowledge delivery into practice: a BONARES case of supporting soil processes modelling

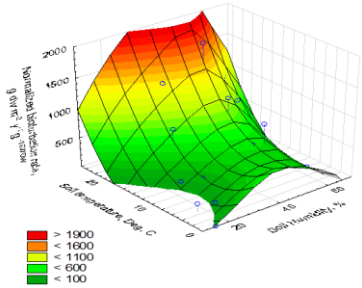
SENCKENBERG
world of biodiversity

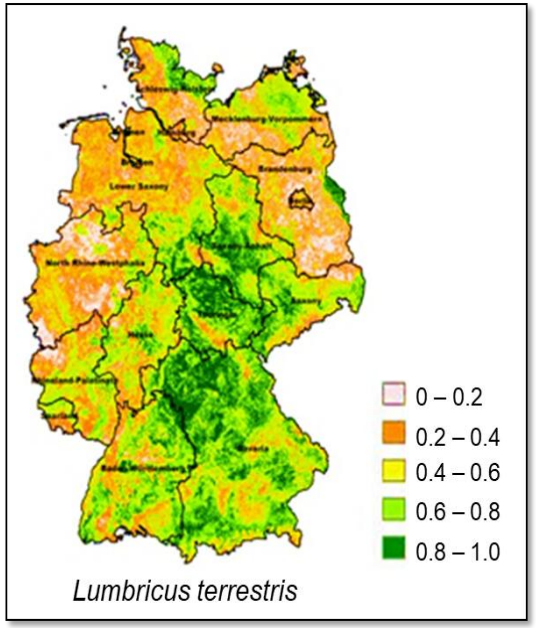
edapho base

External environmental data and BONARES data



Species occurrences



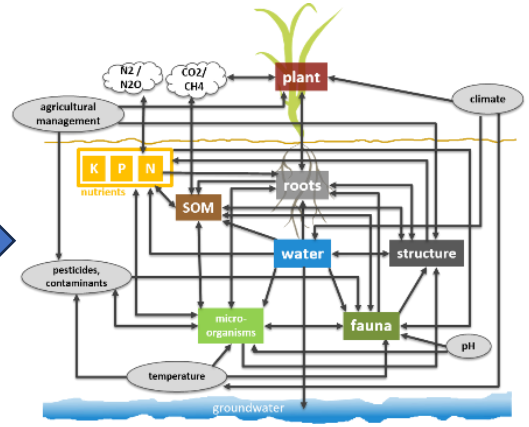




SDM and abundance modelling

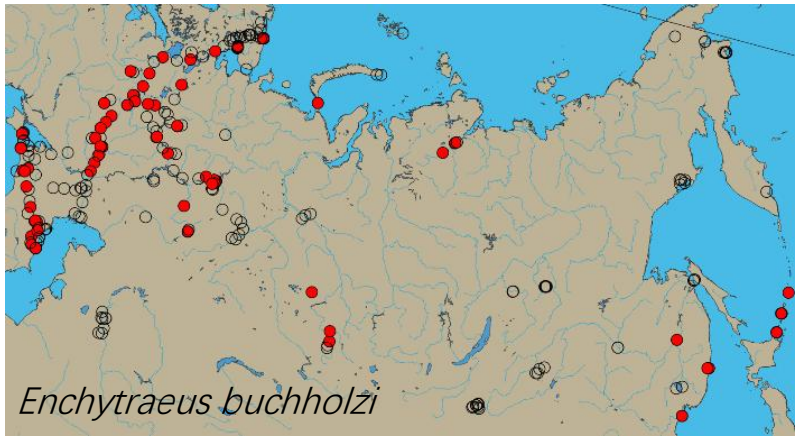
Functional modelling

Processes:
Bioturbation
SOM mobilization and immobilization

BODIUM by BONARES



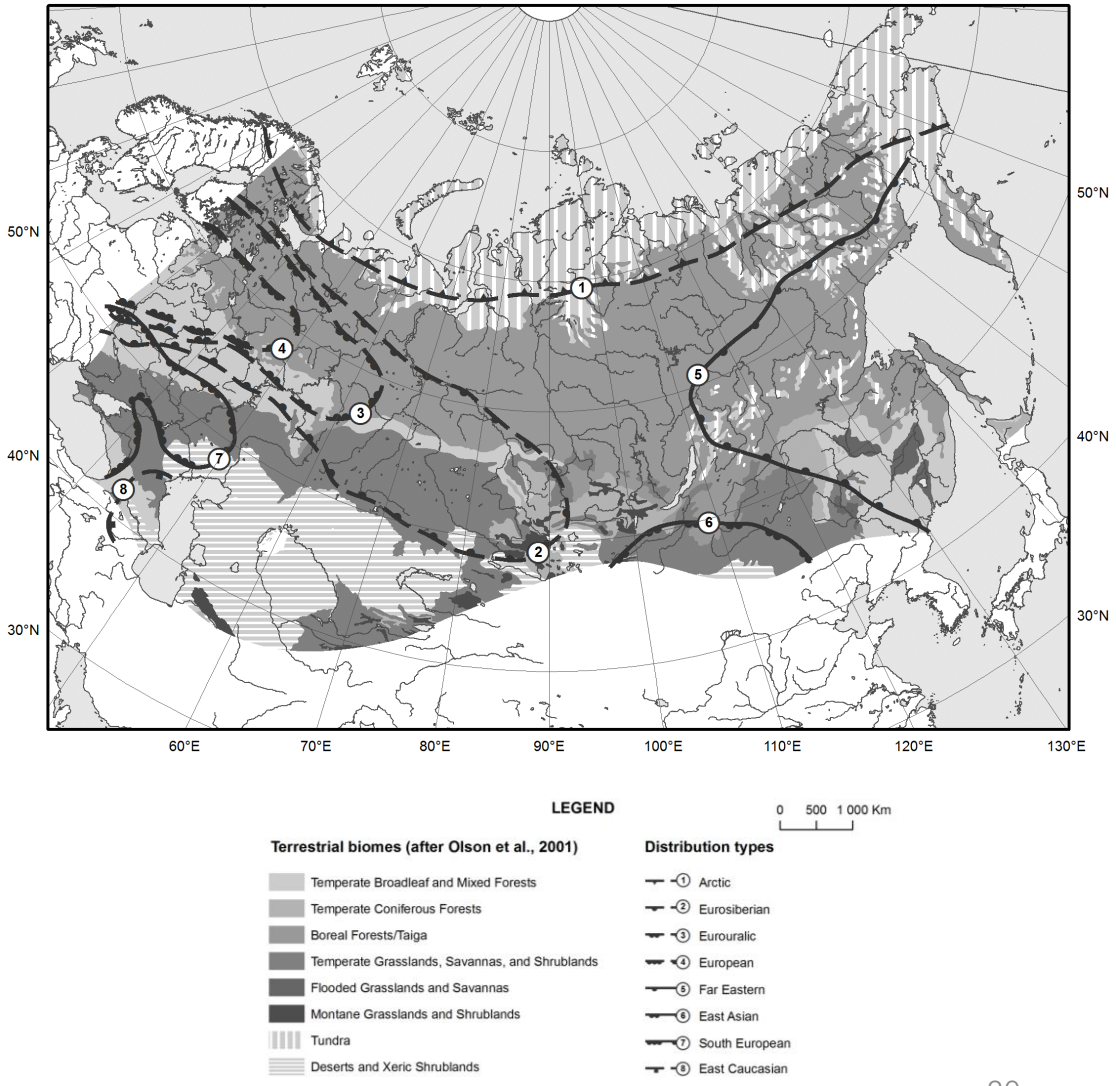
# Usecase II: Soil biogeography and distribution modelling for Enchytraeidae (Oligochaeta)




cost
EUROPEAN COOPERATION IN SCIENCE AND TECHNOLOGY

COST is supported by the EU Framework Programme Horizon 2020

EUdapho base




- Modelling enchytraeid species distribution ranges across northern Eurasia

*Zaitsev et al., Applied Soil Ecology, in revision.*

SENCKENBERG
world of biodiversity

# Systemic targets for the future

- **Fast but reliable** data import process not compromising data quality.
- **Deeper integration within** SGN‘s and external data and knowledge instruments landscape.
- **Outsourcing taxonomic backbone.**
- **Increased visibility** of soil invertebrate collections and related research and promoting SGN‘s data and knowledge through **GBIF** and other international platforms.
- Supporting **spatially explicit** ecological and biogeographic **research**
- Setting new level of standards for **soil biodiversity** data modelling and monitoring soil **biodiversity loss**.
- Bringing solid ground to address **societal needs** related to soil and its functions in light of the functional importance of **soil biodiversity for soil health**.

# Thank you!