Cost Action CA 18237

EUdapho base

Edaphobase Data Structures

Online Training Course 04 September, 2020

David Russell, Stephan Lesch, Sebastian Rick (Senckenberg, Germany)

cost
EUROPEAN COOPERATION
IN SCIENCE & TECHNOLOGY

Funded by the Horizon 2020 Framework Programme
of the European Union

SENCKENBERG
world of biodiversity

SENCKENBERG
world of biodiversity

edapho base

Edaphobase Data Structures

Block I	I. Edaphobase‘s „data philosophy“
	II. Overview of data structures & connections
	Info: Where to find further information
	III. Data on sources
	Info: How personal data is handled
	IV. Data on taxonomy
Block II	V. Data on spatial geography (sites of occurrence)
	VI. Data on soil-biodiversity observations
	Info: Harmonization of quantitative data
	VII. Data on (site) environmental conditions
	Info: Data & Metadata
Block III	VIII. Data on methodology
	IX. For experts: Edaphobase‘s technical data-model

cost
EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

SENCKENBERG
world of biodiversity

edaphobase

I. Edaphobase‘s „data philosophy“

II. Overview of data structures & connections

Info: Where to find further information

III. Data on sources

Info: How personal data is handled

IV. Data on taxonomy

V. Data on spatial geography (sites of occurrence)

VI. Data on soil-biodiversity observations

Info: Harmonization of quantitative data

VII. Data on (site) environmental conditions

Info: Data & Metadata

VIII. Data on methodology

IX. For experts: Edaphobase‘s technical data-model

cost
EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

All databases are structured according to their **goals**

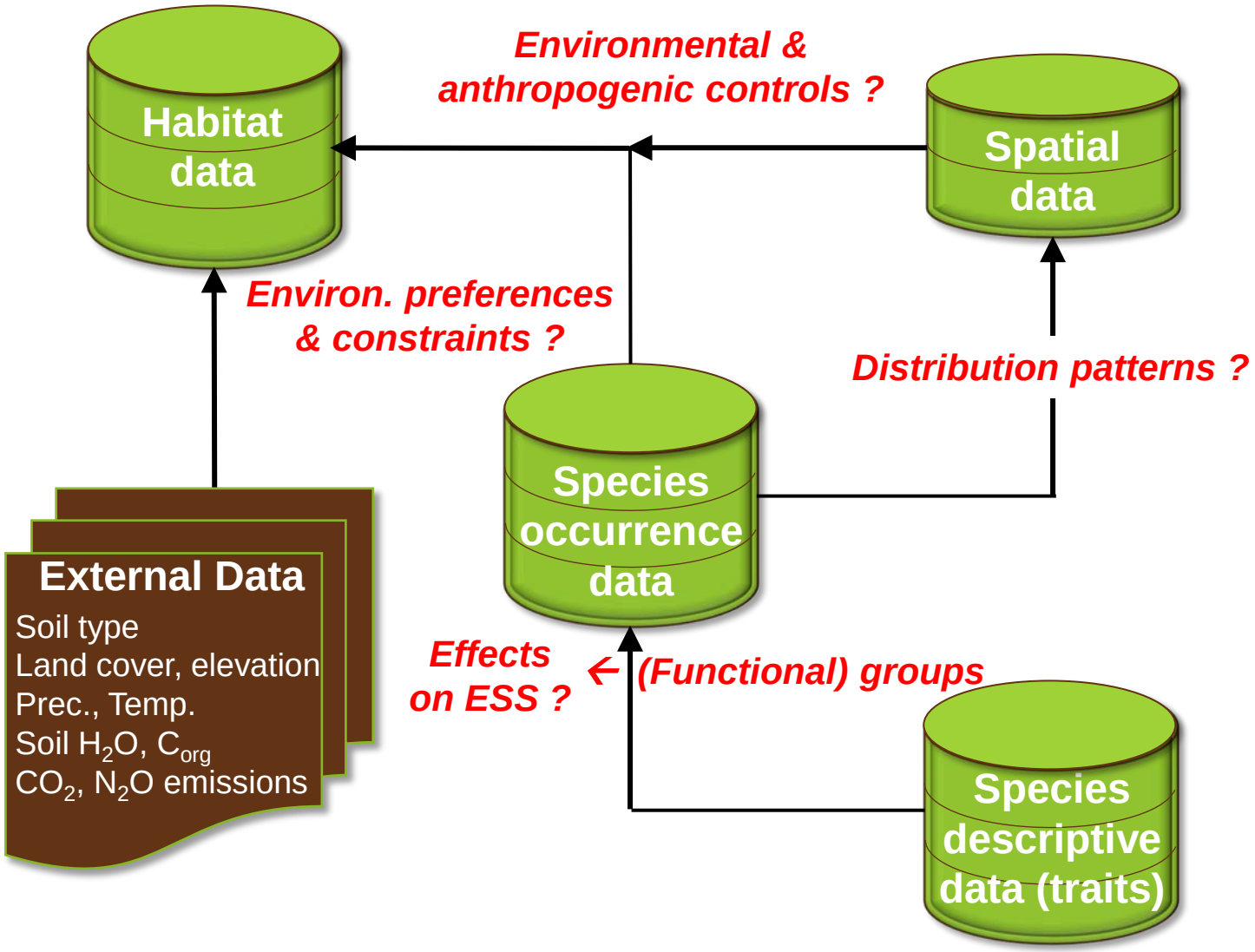
especially: how the data is to be (re-)used
(= „*requirements*")

Edaphobase explicitly developed to understand:

Soil biodiversity ecology

- **Species distribution (and its drivers)**
- **Species autecology (niche space)**
- **Soil-biodiversity of areas or habitat types**
- ***Soil-biodiversity functional roles***

Intended answerable questions with Edaphobase:



Further Edaphobase intentions:

- **Integrate heterogeneous sources**
- **Accommodate diverse (data-provider) purposes**
  - Data repository (→ DOIs)
  - Research data management (RDM)

    Biodiversity surveys & monitoring programmes
    Research projects
    Ecotoxicology
    Experiments
    …
  - Collection management
- **Structure & harmonize to allow common data re-use**
- **Unsure high data quality**
- **Follow FAIR principles**

Edaphobase ≠ Data Repository
(storing „stand alone“ tables of data)

Edaphobase = a <u>Data Warehouse</u>
(integrating, harmonizing and storing all data together)

(Definition á Kimball & Ross 2002, Inmon 2005):

- *Subject-oriented; Domain specific (= soil biodiversity)*
- *Integrate and structure heterogeneous data (from multiple sources)*
- *Data homogeneously structured for common query and analysis*
- *Time-variant; i.e., data from different acquisition dates together*
- *Non-volatile; i.e., once data is stored, it should not be altered*
  → *possible corrections*

SENCKENBERG
world of biodiversity

edapho base

I. Edaphobase‘s „data philosophy“

II. Overview of data structures & connections

Info: Where to find further information

III. Data on sources

Info: How personal data is handled

IV. Data on taxonomy

V. Data on spatial geography (sites of occurrence)

VI. Data on soil-biodiversity observations

Info: Harmonization of quantitative data

VII. Data on (site) environmental conditions

Info: Data & Metadata

VIII. Data on methodology

IX. For experts: Edaphobase‘s technical data-model

cost
EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

I. Edaphobase‘s „data philosophy“

II. Overview of data structures & connections

Info: Where to find further information

III. Data on sources

> 600 data fields (variables)

= options !

Most *not* required (many rarely used)

Edaphobase = complex

Info: Data & Metadata

VIII. Data on methodology

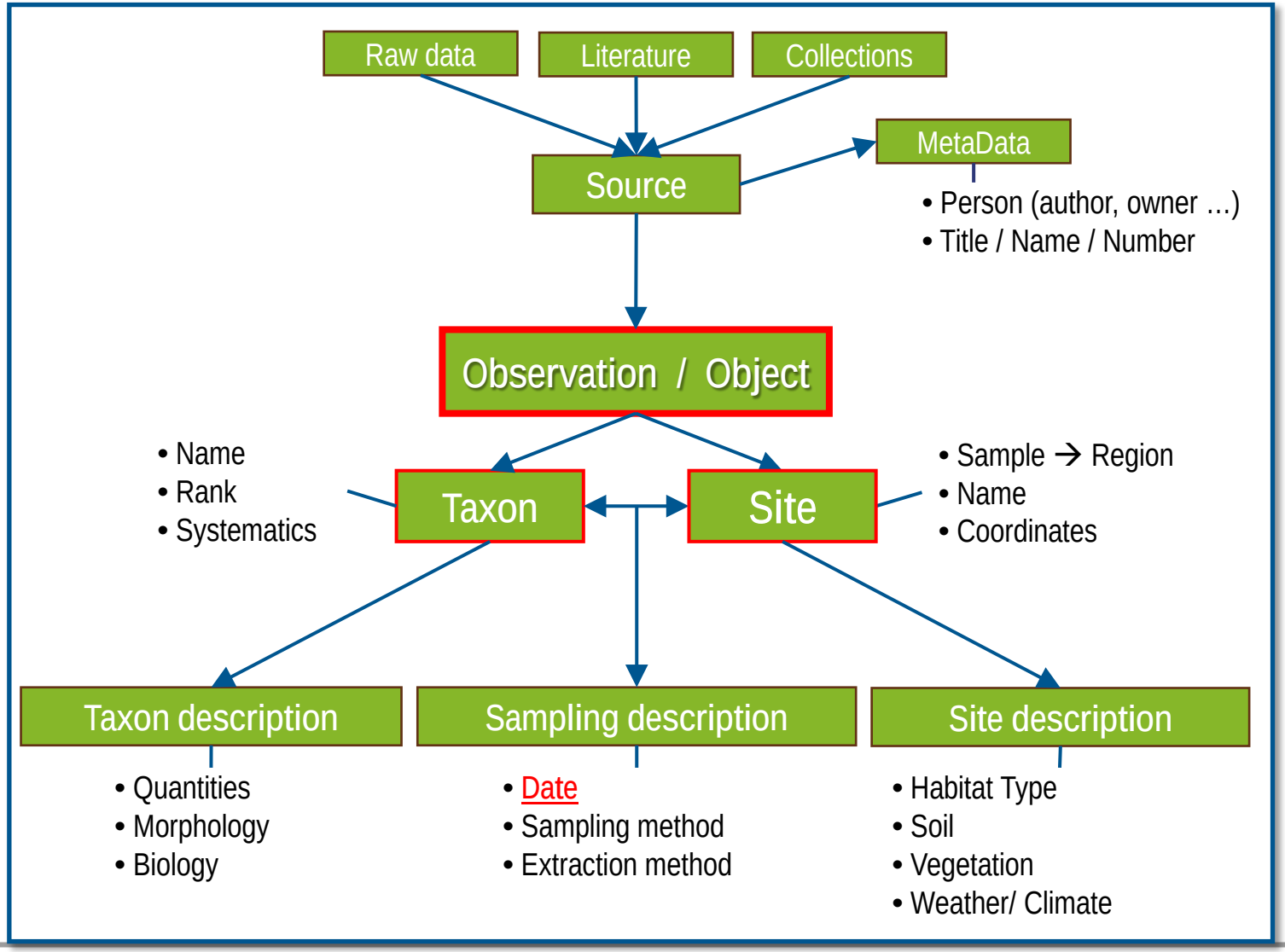
IX. For experts: Edaphobase‘s technical data-model

Edaphobase data model structured to assess:

What	→	Which taxa & their quantities
Where	→	Which sites of occurrence
When	→	Sampling dates/periods
Which conditions	→	Site environmental (meta)data
Who	→	Who sampled; Data owners
How	→	Methodological metadata

Edaphobase Overview

General data-model concept



Edaphobase Overview

Common data sheets & Edaphobase data linkage

SITE (Spatial / Habitat data)

Biogeo-region	Country	Site	Plot	Sample	Lat	Long	Soil …	…	Habi-tat…	…	Veg./ Land Use…	…
-	-	-			-	-	-	-	-	-	-	-
-	-	-			-	-	-	-	-	-	-	-
-	-	-			-	-	-	-	-	-	-	-
					-	-	-	-	-	-	-	-

via Site

TAXON (Species occurrence data)

Species	Site	Sampling date	Abund.	Biomass*	Length*	Width*
-	-	-	-	-	-	-
-			-	-	-	-
-		-	-	-	-	-
-			-	-	-	-
-	-	-	-	-	-	-
-			-	-	-	-
-		-	-	-	-	-
-			-	-	-	-

* Measured; spp overall averages in descriptive data

Edaphobase Overview

Common data sheets & Edaphobase data linkage

SITE *(Spatial / Habitat data)*

Biogeo-region	Country	Site	Plot	Sample	Lat	Long	Soil …	…	Habi-tat…	…	Veg./ Land Use…	…
-	-	-			-	-	-	-	-	-	-	-
-	-	-			-	-	-	-	-	-	-	-
-	-	-			-	-	-	-	-	-	-	-
					-	-						-

via Site

(+ Date)

= Observation

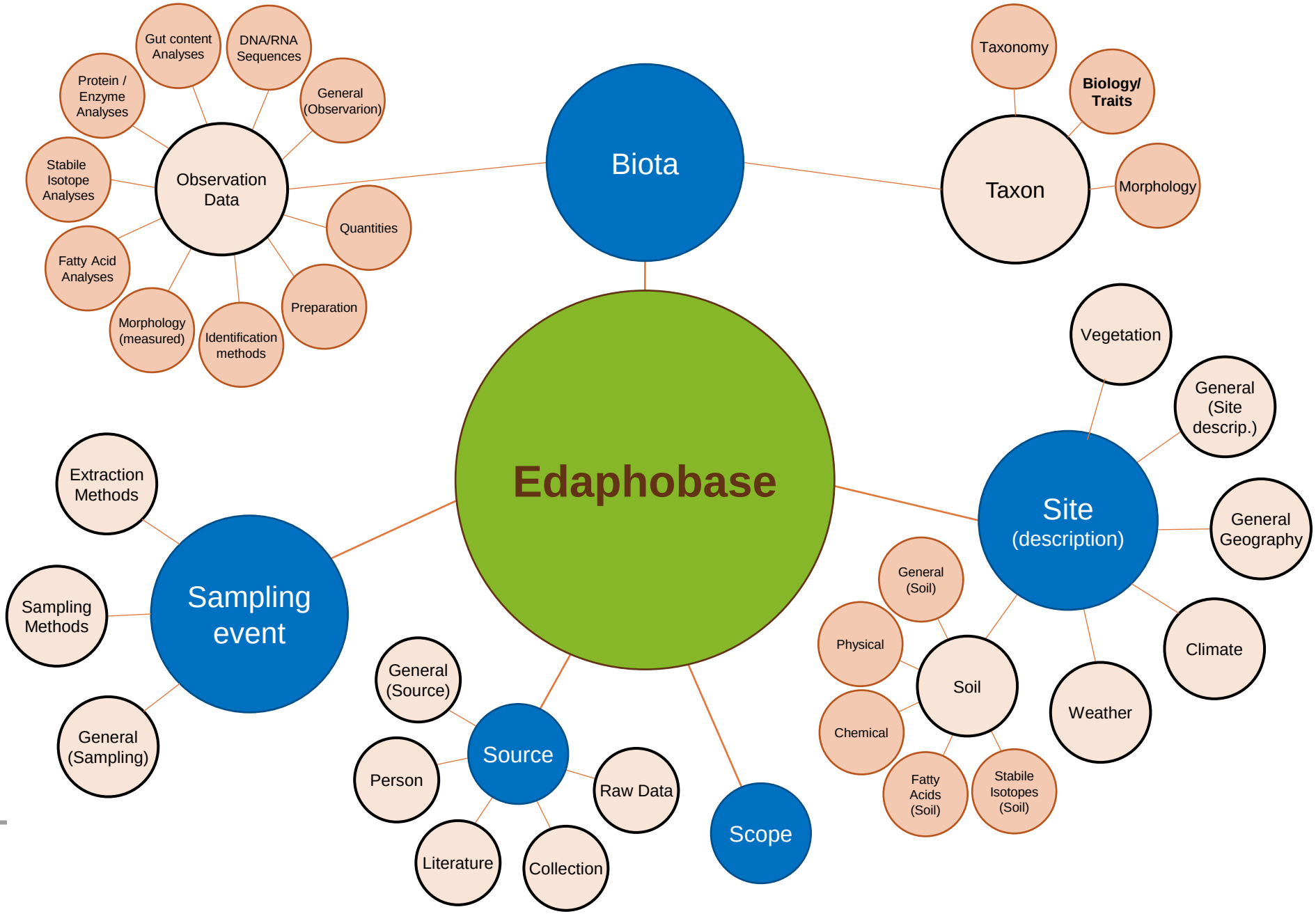
TAXON

(Species occurrence data)

Species	Site	Sampling date	Abund.	Biomass*	Length*	Width*
-	-	-	-	-	-	-
-			-	-	-	-
-			-	-	-	-
-			-	-	-	-
-	-	-	-	-	-	-
-			-	-	-	-
-			-	-	-	-
-			-	-	-	-

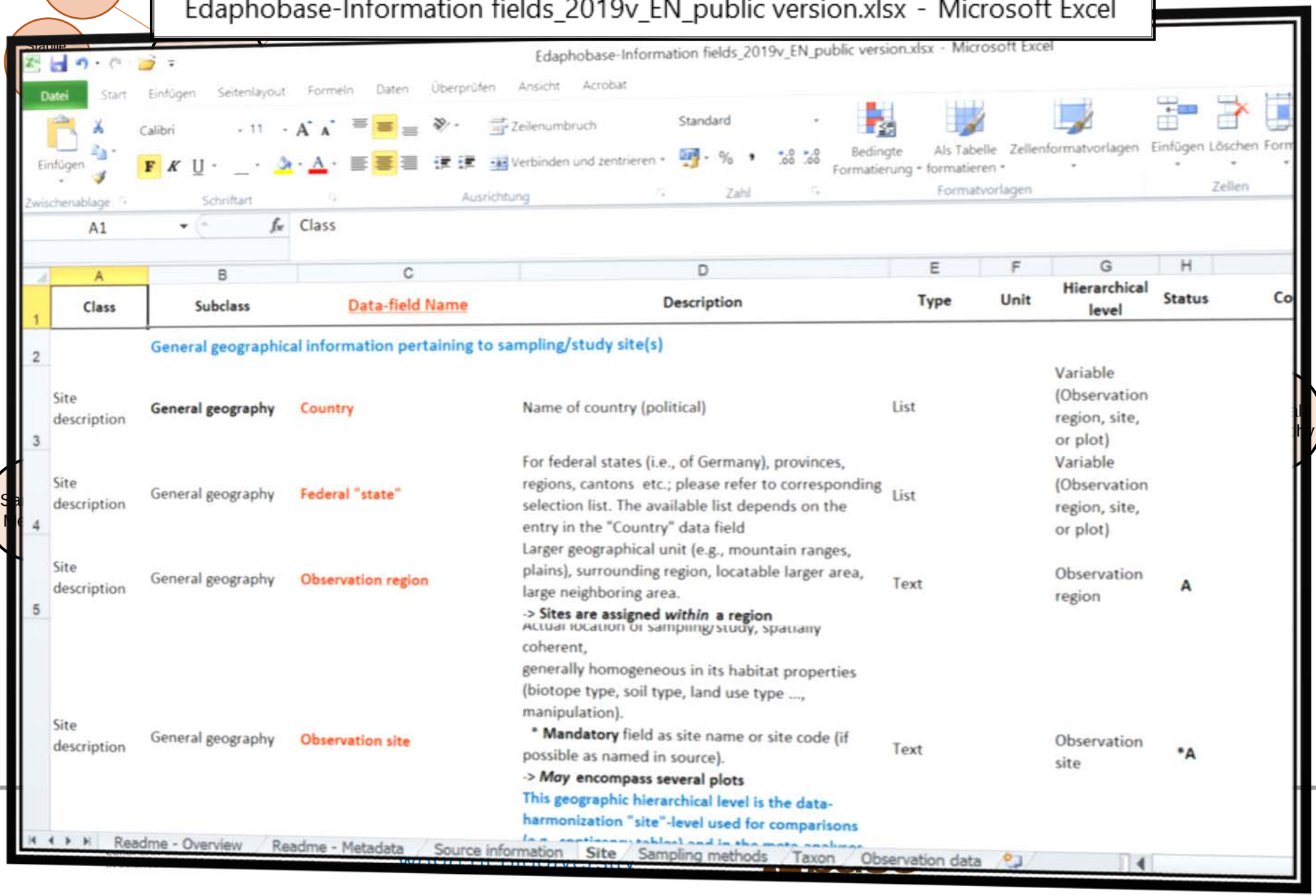
* Measured; spp overall averages in descriptive data

Edaphobase data structures and linkages



Further Info | → **edaphobase.org** // Nextcloud / Edaphobase – Information fields (variables)

Edaphobase-Information fields_2019v_EN_public version.xlsx - Microsoft Excel



Edaphobase-Information fields_2019v_EN_public version.xlsx - Microsoft Excel

	A	B	C	D	E	F	G	H
1	Class	Subclass	Data-field Name	Description	Type	Unit	Hierarchical level	Status
2		General geographical information pertaining to sampling/study site(s)						
3	Site description	General geography	Country	Name of country (political)	List		Variable (Observation region, site, or plot)	
4	Site description	General geography	Federal "state"	For federal states (i.e., of Germany), provinces, regions, cantons etc.; please refer to corresponding selection list. The available list depends on the entry in the "Country" data field	List		Variable (Observation region, site, or plot)	
5	Site description	General geography	Observation region	Larger geographical unit (e.g., mountain ranges, plains), surrounding region, locatable larger area, large neighboring area. -> **Sites are assigned *within* a region**	Text		Observation region	A
	Site description	General geography	Observation site	Actual location of sampling/study, spatially coherent, generally homogeneous in its habitat properties (biotope type, soil type, land use type ..., manipulation). * **Mandatory** field as site name or site code (if possible as named in source). -> ***May* encompass several plots** This geographic hierarchical level is the data-harmonization "site"-level used for comparisons	Text		Observation site	*A

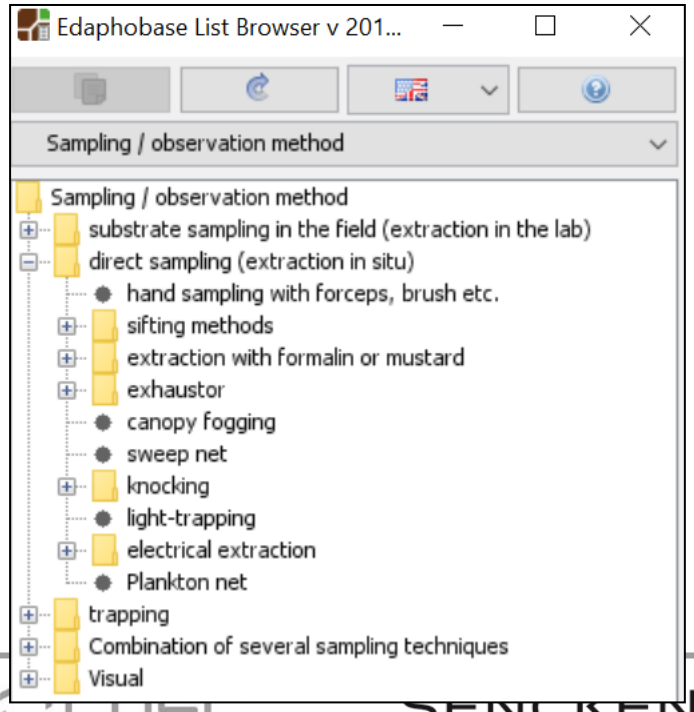
Readme - Overview | Readme - Metadata | Source information | Site | Sampling methods | Taxon | Observation data

Categorical variables

(i.e. Taxonomy, Habitat types, Soil types, Soil Horizon, etc.)

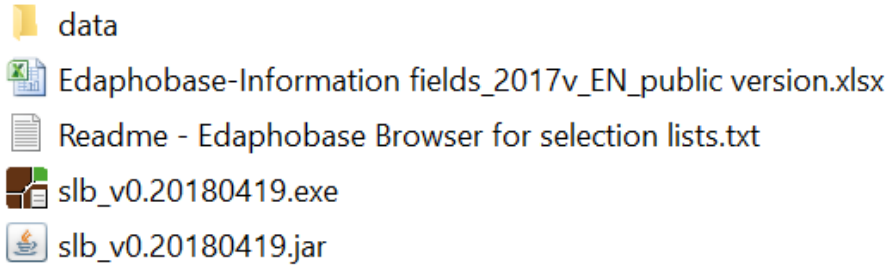
→ Standardized vocabularies („authority lists“)
(data comparability, avoids typo errors)

„Edaphobase selection-list browser“
slb_v0.20180419.exe



Further Info

- edaphobase.org
- Nextcloud \
  Edaphobase - Information Fields (variables) \
  Authority List Browser



Categorical variables

(i.e. Taxonomy, Habitat types, Soil types, Soil horizon, etc.)

→ Standardized vocabularies („authority lists“)
(data comparability, avoids typo errors)

Use international standards

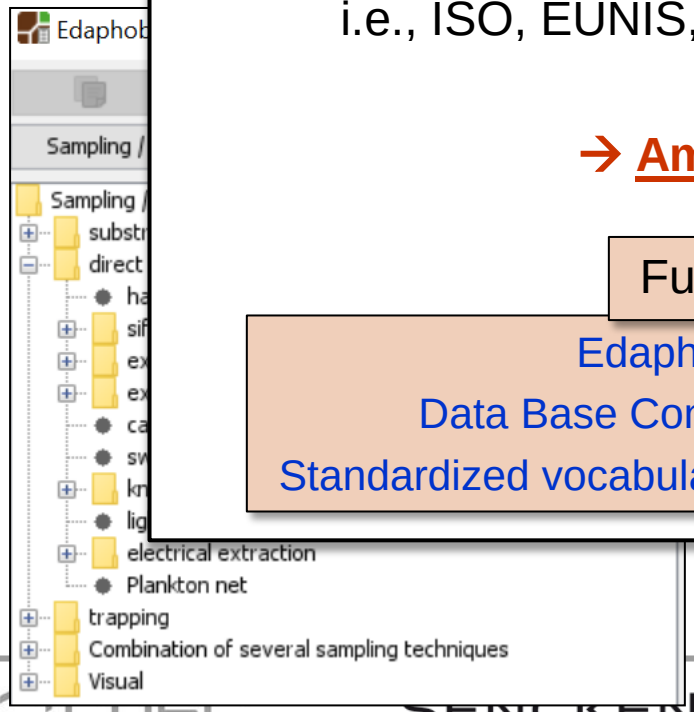
(where available)

i.e., ISO, EUNIS, CORINE, WRB etc.

→ **Amendable**

Further Info

Edaphobase.org /
Data Base Content and Structure /
Standardized vocabulary for categorical variables



ion Fields (variables)\

7v_EN_public version.xlsx

selection lists.txt

slb_v0.20180419.exe

slb_v0.20180419.jar

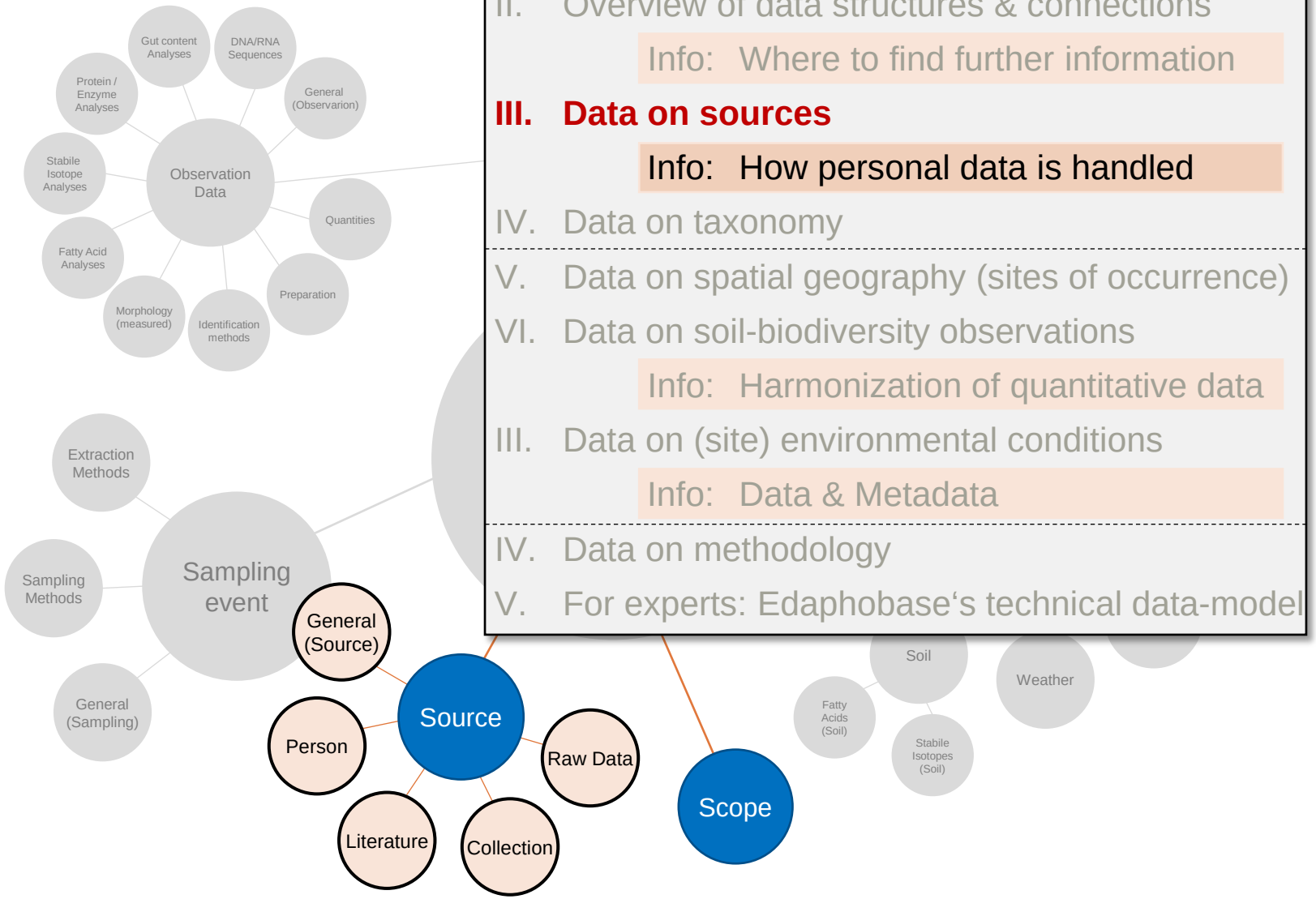
Database Vocabulary!

(so we're all on the same page today)

Data set

Data record

	A	B	C	D	E	F	G	
1	Taxon	Count in sample(s)	Country	Observation area	Observation site	Latitude	Longitude	
2	Oniscus asellus Linnaeus, 1758	1	Germany	NP Eifel	Schlitterley (EB_NW33)	50.6203	6.4938	Eichen
3	Ligidium hypnorum (Cuvier, 1792)	1	Germany	NP Eifel	Naturwald Nord (EB_NW23)	50.5698	6.3601	boden
4	Trichoniscus	1	Germany	NP Eifel	Langerscheid 2 (EB_NW22)	50.5461	6.3405	boden
5	Oniscus asellus Linnaeus, 1758	1	Germany	NP Eifel	Langerscheid 1 (EB_NW13)	50.5500	6.3415	Fichter
6	Trichoniscus	1	Germany	NP Eifel	Langerscheid 2 (EB_NW22)	50.5461	6.3405	boden
7	Oniscus asellus Linnaeus, 1758	1	Germany	NP Eifel	Schlitterley (EB_NW33)	50.6203	6.4938	Eichen
8	Ligidium hypnorum (Cuvier, 1792)	1	Germany	NP Eifel	Langerscheid 1 (EB_NW13)	50.5500	6.3415	Fichter
9	Porcellio scaber Latreille, 1804	1	Germany	NP Eifel	Weiersheld (EB_NW32)	50.6263	6.4154	Eichen
10	Trichoniscus	1	Germany	NP Eifel	Naturwald Nord (EB_NW23)	50.5698	6.3601	boden
11	Porcellio scaber Latreille, 1804	1	Germany	NP Eifel	Schlitterley (EB_NW33)	50.6203	6.4938	Eichen
12	Oniscus asellus Linnaeus, 1758	1	Germany	NP Eifel	Langerscheid 1 (EB_NW13)	50.5500	6.3415	Fichter
13	Philoscia	1	Germany	NP Eifel	Weiersheld (EB_NW32)	50.6263	6.4154	Eichen
14	Oniscus asellus Linnaeus, 1758	1	Germany	NP Eifel	Weiersheld (EB_NW32)	50.6263	6.4154	Eichen
15	Oniscus asellus Linnaeus, 1758	1	Germany	NP Eifel	Langerscheid 1 (EB_NW13)	50.5500	6.3415	Fichter
16	Oniscus asellus Linnaeus, 1758	1	Germany	NP Eifel	Schlitterley (EB_NW33)	50.6203	6.4938	Eichen
17	Oniscus asellus Linnaeus, 1758	1	Germany	NP Eifel	Langerscheid 1 (EB_NW13)	50.5500	6.3415	Fichter
18	Oniscus asellus Linnaeus, 1758	2	Germany	NP Eifel	B265 Gemünd (EB_NW31)	50.5803	6.5020	Eichen
19	Oniscus asellus Linnaeus, 1758	1	Germany	NP Eifel	Wächterbuche (EB_NW21)	50.5402	6.3382	boden
20	Oniscus asellus Linnaeus, 1758	1	Germany	NP Eifel	Schlitterley (EB_NW33)	50.6203	6.4938	Eichen
21	Oniscus asellus Linnaeus, 1758	1	Germany	NP Eifel	B265 Gemünd (EB_NW31)	50.5803	6.5020	Eichen
22	Philoscia	1	Germany	NP Eifel	B265 Gemünd (EB_NW31)	50.5803	6.5020	Eichen
23	Oniscus asellus Linnaeus, 1758	1	Germany	NP Eifel	Wächterbuche (EB_NW21)	50.5402	6.3382	boden
24	Oniscus asellus Linnaeus, 1758	1	Germany	NP Eifel	Naturwald Nord (EB_NW23)	50.5698	6.3601	boden


I. Edaphobase‘s „data philosophy“
II. Overview of data structures & connections
Info: Where to find further information
III. Data on sources
Info: How personal data is handled
IV. Data on taxonomy
V. Data on spatial geography (sites of occurrence)
VI. Data on soil-biodiversity observations
Info: Harmonization of quantitative data
III. Data on (site) environmental conditions
Info: Data & Metadata
IV. Data on methodology
V. For experts: Edaphobase‘s technical data-model
Observation Data
Gut content Analyses
DNA/RNA Sequences
General (Observarion)
Protein / Enzyme Analyses
Stabile Isotope Analyses
Quantities
Fatty Acid Analyses
Preparation
Morphology (measured)
Identification methods
Sampling event
Extraction Methods
Sampling Methods
General (Sampling)
Source
General (Source)
Person
Raw Data
Literature
Collection
Scope
Soil
Weather
Fatty Acids (Soil)
Stabile Isotopes (Soil)

Edaphobase Data Sources

Raw data

	valid taxon	country	district	Habitat type
				658 / 721
237,459	Ceratophysella ...	France (FR)	Midi-Pyrénées (FR-N)	Deciduous and mixed wo
228,064	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Deciduous and mixed wo
228,139	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Deciduous and mixed wo
228,163	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Deciduous and mixed wo
228,168	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Deciduous and mixed wo
228,208	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Deciduous and mixed wo
228,226	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Deciduous and mixed wo
228,246	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Deciduous and mixed wo
228,292	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Deciduous and mixed wo
228,545	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Deciduous and mixed wo
228,624	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Fir-spruce plantations (4
228,650	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Fir-spruce plantations (4
228,672	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Fir-spruce plantations (4
228,738	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Deciduous and mixed wo
228,757	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Deciduous and mixed wo
231,202	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Riparian and forest fring
231,293	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Riparian and forest fring
232,143	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Deciduous and mixed wo
232,160	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Deciduous and mixed wo
233,258	Ceratophysella ...	Czech Republic ...	South Moravian Region (Jihomoravský kra...	Natural dry grasslands a
233,259	Ceratophysella ...	Czech Republic ...	South Moravian Region (Jihomoravský kra...	Natural dry grasslands a
233,330	Ceratophysella ...	Czech Republic ...	South Moravian Region (Jihomoravský kra...	Natural dry grasslands a

- Research projects
- Monitoring results
- Student courses
- Personal observations

Literature

- Journal articles
- Books
- Reports
- „Grey“ literature

Collections

- Museum collections
- Sub-collections
- Collection objects
- Types

Edaphobase Data Sources

Raw data

	valid taxon	country	district	Habitat type
				658 / 721
237,459	Ceratophysella ...	France (FR)	Midi-Pyrénées (FR-N)	Deciduous and mixed wo
228,064	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Deciduous and mixed wo
228,139	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Deciduous and mixed wo
228,163	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Deciduous and mixed wo
228,168	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Deciduous and mixed wo
228,208	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Deciduous and mixed wo
228,226	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Deciduous and mixed wo
228,246	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Deciduous and mixed wo
228,292	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Deciduous and mixed wo
228,545	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Deciduous and mixed wo
228,624	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Fir-spruce plantations (4
228,650	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Fir-spruce plantations (4
228,672	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Fir-spruce plantations (4
228,738	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Deciduous and mixed wo
228,757	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Deciduous and mixed wo
231,202	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Riparian and forest fring
231,293	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Riparian and forest fring
232,143	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Deciduous and mixed wo
232,160	Ceratophysella ...	Germany (DE)	Baden-Württemberg (DE-BW)	Deciduous and mixed wo
233,258	Ceratophysella ...	Czech Republic ...	South Moravian Region (Jihomoravský kra...	Natural dry grasslands a
233,259	Ceratophysella ...	Czech Republic ...	South Moravian Region (Jihomoravský kra...	Natural dry grasslands a
233,330	Ceratophysella ...	Czech Republic ...	South Moravian Region (Jihomoravský kra...	Natural dry grasslands a

- Dataset name
- Project title
- Subproject
- PI / Data Owner
- …

Literature

- Article / Chapter title
- Journal / Book
- Pages
- Authors
- …

Collections

- Collection name
- Sub-collection name
- Collection object nr.
- Type (Y/N)
- …

Edaphobase Data Sources

Raw data

Further Info

Collections

Personal Information:

- Name (first name, family name)
- Title
- Institution
- Address
- E-Mail address
- (Phone number)

→ Only „Name“ is shown online! (to cite data source!)

→ Edaphobase does **not** share with **any** external source!

- Datas
- Proje
- Subproject
- PI / Data Owner
- …

- Article / Chapter title
- Pages
- …

- on name
- ection name
- Collection object nr.
- Type (Y/N)
- …

Sampling Scope

(Background information on purpose of data set)

➢ **Needed for data (re-)use**
(filtering appropriate data in analysis tools)

i.e., *not* appropriate:

- single observations as site representative
- species-level studies in community-level analyses
- presence/absence data in quantitative regressions

3 (binary) variables:

Sampling Effort	→	*single sample / multiple samples*
Quantification level	→	*qualitative / quantitative*
Species composition level	→	*single species / community (spp. composition)*

I. Edaphobase‘s „data philosophy“

II. Overview of data structures & connections

Info: Where to find further information

III. Data on sources

Info: How personal data is handled

IV. Data on taxonomy

V. Data on spatial geography (sites of occurrence)

VI. Data on soil-biodiversity observations

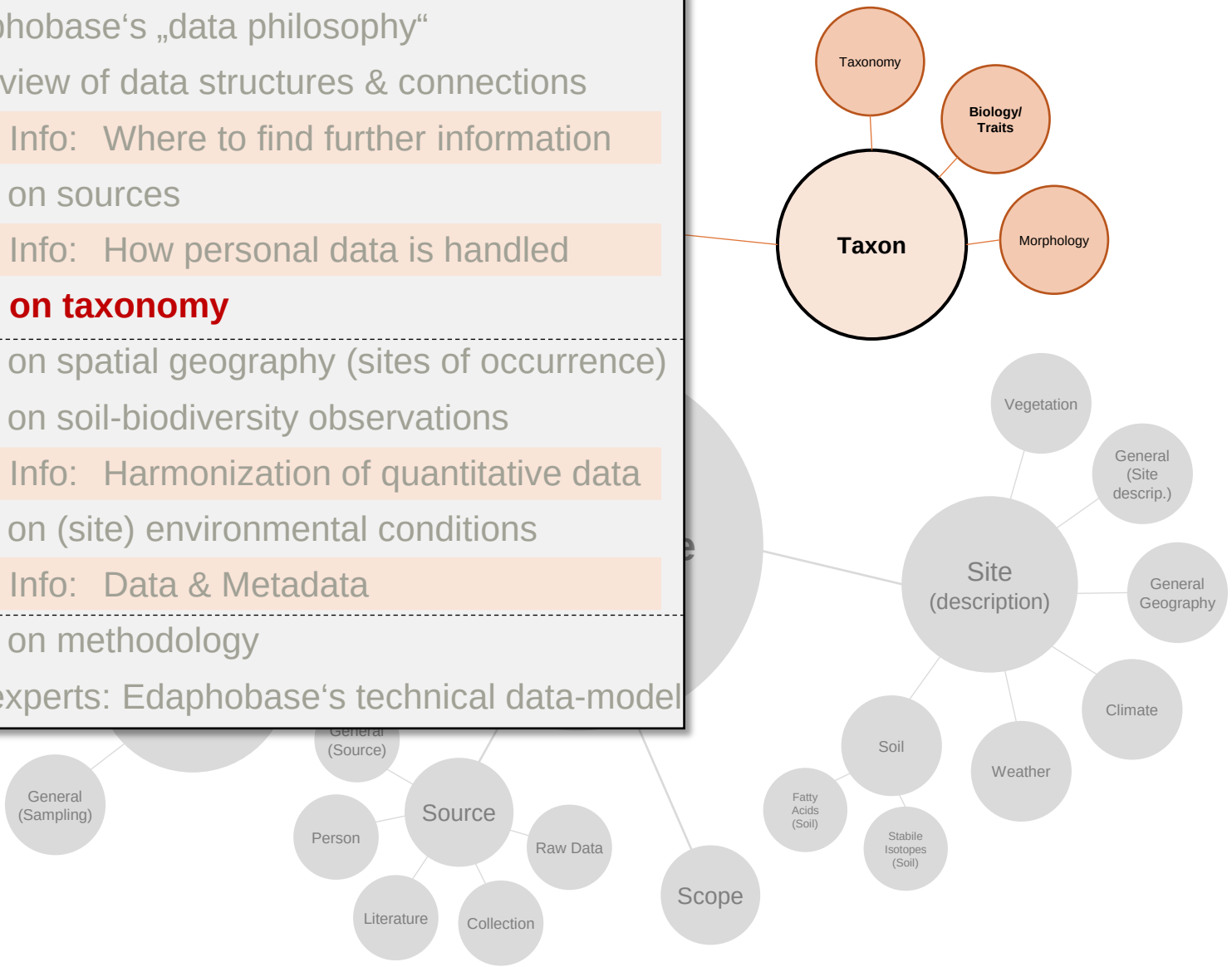
Info: Harmonization of quantitative data

VII. Data on (site) environmental conditions

Info: Data & Metadata

VIII. Data on methodology

IX. For experts: Edaphobase‘s technical data-model



Edaphobase Taxonomic "Backbone"

Only full taxonomic nomenclature = distinct & unambiguous
(data comparability!)

Includes:

- taxon name (i.e., genus, species epithet)
- describing author(s)
- year of description
- brackets or not

➢ follows int'l nomenclatural rules

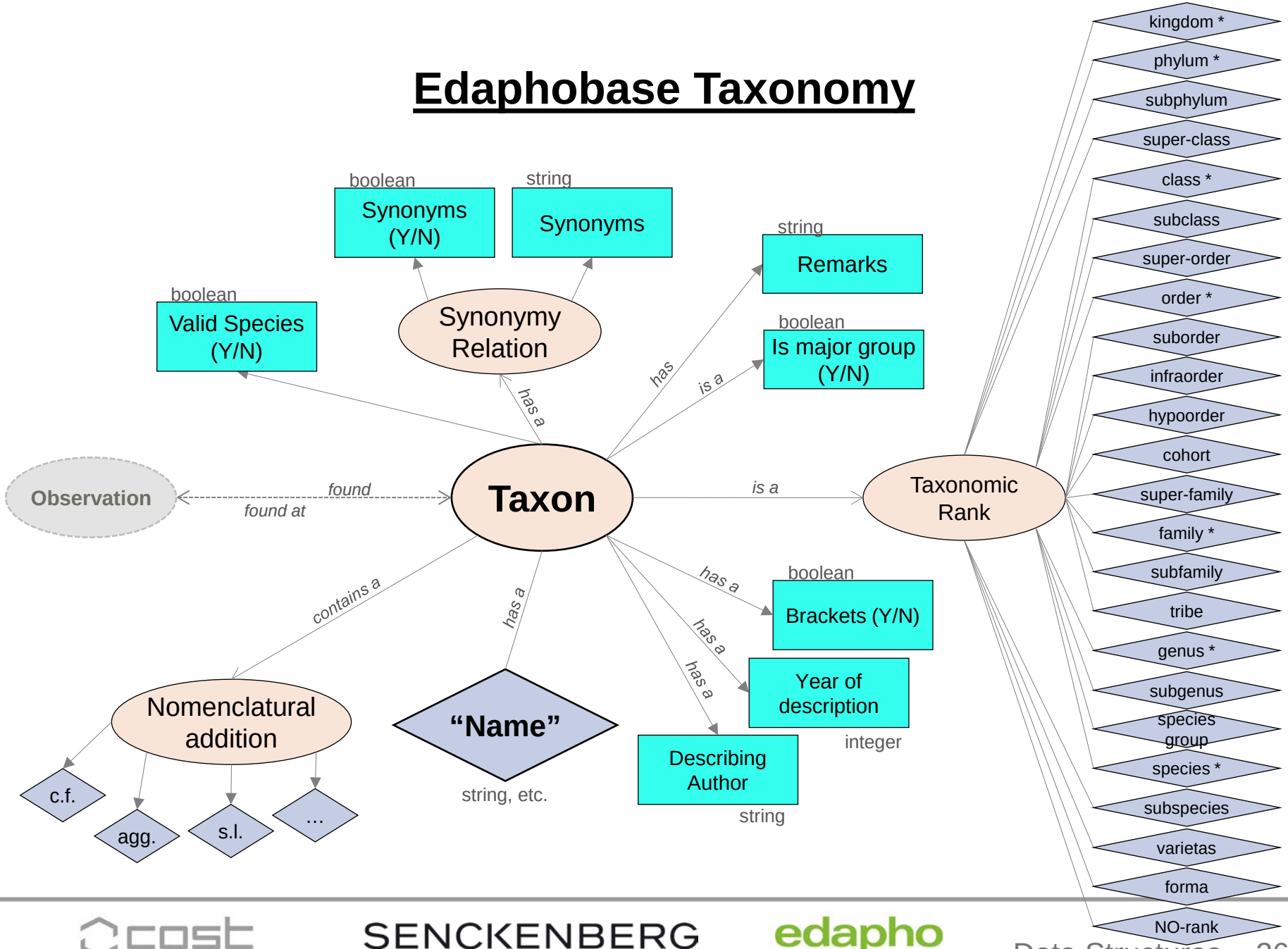
Edaphobase maintains a taxonomic „backbone" for all included organism groups

- Based on authoritative sources (i.e., Fauna Europaea, collembola.org, int'l experts)
- As up to date as possible
- Avoids typographic errors *(data comparability!)*

➢ Includes synonym relationships
➢ Amendable! (new taxa can be added as needed)

→ Taxonomic review boards needed!

Edaphobase Taxonomy



Edaphobase Taxonomy → Traits

Work in Progress

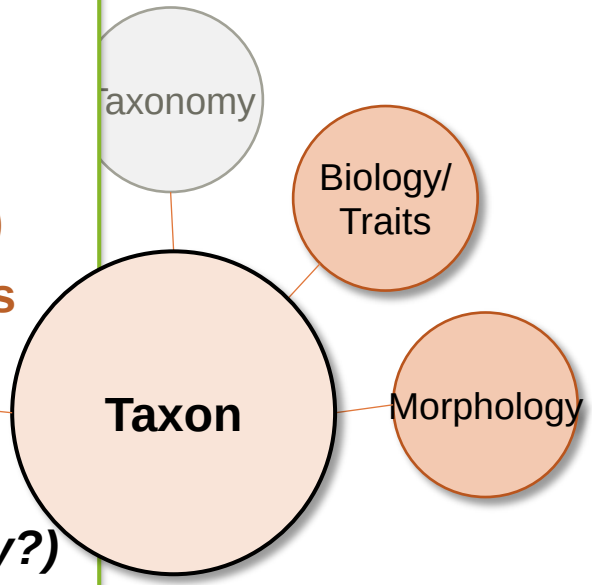
Currently very rudimentary

- **Trophic level**
- **Nematode feeding types**
- **Life form types (earthworm, etc.)**
- **(many) morphological characters**
- **…**

Plan → *strongly expand!*

(links to, e.g., BETSI, EcoTaxonomy?)

- ***Behavioural / Physiological traits***
  - ***Effect traits***
  - ***Response Traits***
- ***Morphological characters***



Edaphobase Taxonomy → Traits

Work in Progress

Currently very rudimentary

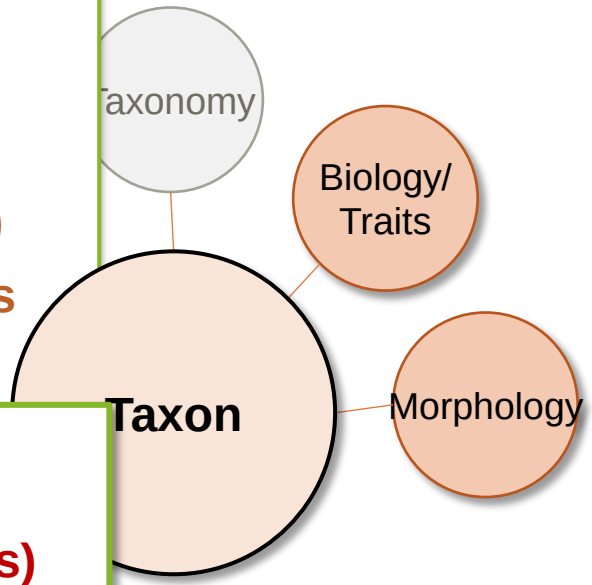
- **Trophic level**
- **Nematode feeding types**
- **Life form types (earthworm, etc.)**
- **(many) morphological characters**
- **…**

(lin

- ➢ ***Morphological characters***

Linked to the Taxon !
(not individual measurements)

Linked to any systematic level



Edaphobase Overview

Common data sheets & Edaphobase data linkage

SITE (Spatial / Habitat data)

Biogeo-region	Country	Site	Plot	Sample	Lat	Long	Soil …	…	Habi-tat…	…	Veg./ Land Use…	…
-	-	-			-	-	-	-	-	-	-	-
-	-	-			-	-	-	-	-	-	-	-
-	-	-			-	-	-	-	-	-	-	-
					-	-	-	-	-	-	-	-

via Site

TAXON (Species occurrence data)

Species	Site	Sampling date	Abund.	Biomass*	Length*	Width*
-	-	-	-	-	-	-
-			-	-	-	-
-			-	-	-	-
-			-	-	-	-
-	-	-	-	-	-	-
-			-	-	-	-
-			-	-	-	-
-			-	-	-	-

* Measured; spp overall averages in descriptive data

SITE (Spatial / Habitat data)

Biogeo-region	Country	Site	Plot	Sample	Lat	Long	Soil …	…	Habi-tat…	…	Veg./ Land Use…	…
-	-	-			-	-	-	-	-	-	-	-
-	-	-			-	-	-	-	-	-	-	-
-	-	-			-	-	-	-	-	-	-	-
							-	-	-	-	-	-

Site

TAXON (Species occurrence data)

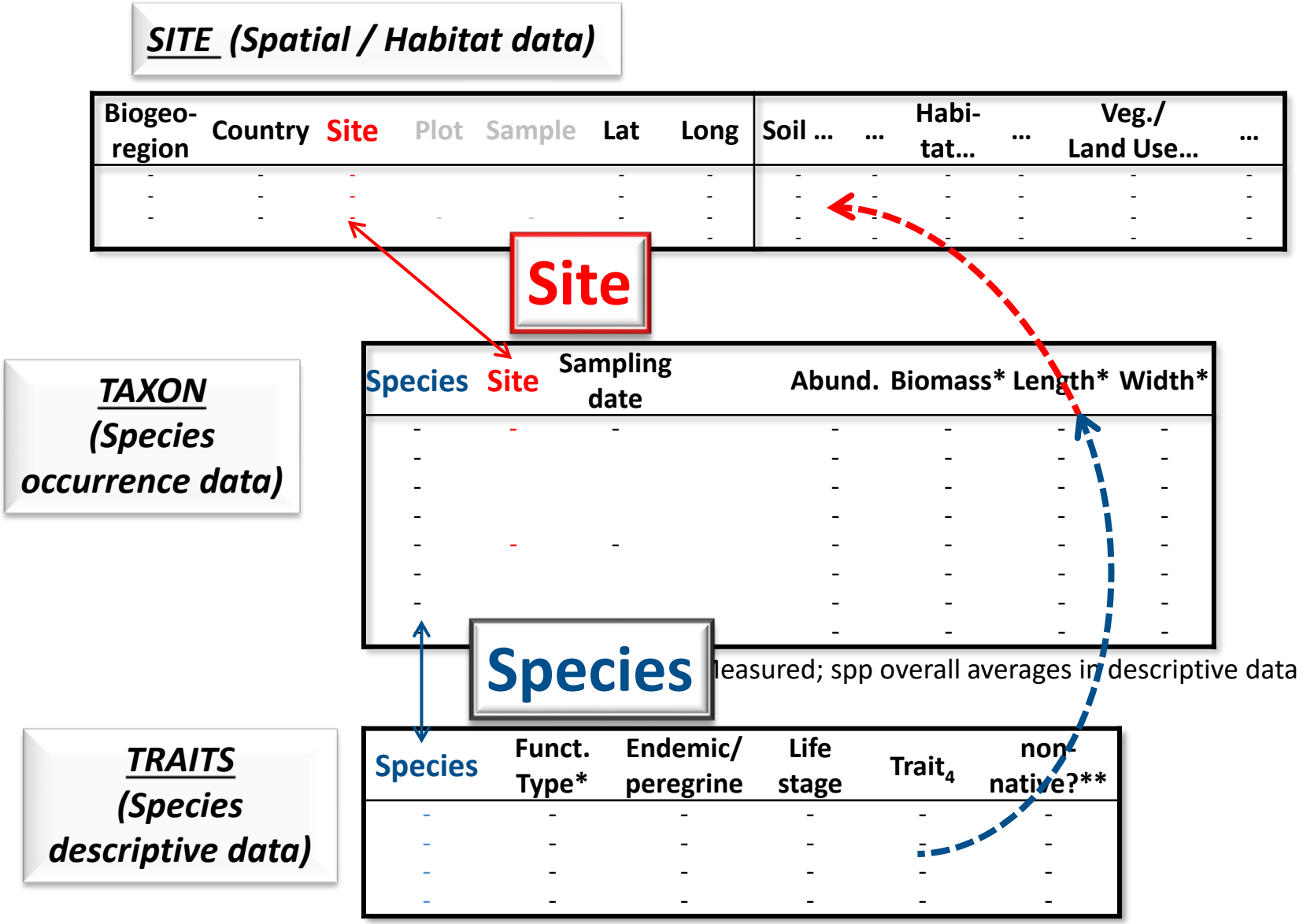
Species	Site	Sampling date	Abund.	Biomass*	Length*	Width*
-	-	-	-	-	-	-
-			-	-	-	-
-			-	-	-	-
-			-	-	-	-
-	-	-	-	-	-	-
-			-	-	-	-
-			-	-	-	-
			-	-	-	-

Species

easured; spp overall averages in descriptive data

TRAITS (Species descriptive data)

Species	Funct. Type*	Endemic/ peregrine	Life stage	$Trait_4$	non-native?**
-	-	-	-	-	-
-	-	-	-	-	-
-	-	-	-	-	-
-	-	-	-	-	-


SITE (Spatial / Habitat data)
Biogeo-region
Country
Site
Plot
Sample
Lat
Long
Soil ...
...
Habi-tat...
...
Veg./ Land Use...
...
Site
TAXON (Species occurrence data)
Species
Site
Sampling date
Abund.
Biomass*
Length*
Width*
Species
measured; spp overall averages in descriptive data
TRAITS (Species descriptive data)
Species
Funct. Type*
Endemic/ peregrine
Life stage
Trait4
non-native?**

I. Edaphobase‘s „data philosophy“

II. Overview of data structures & connections

Info: Where to find further information

III. Data on sources

Info: How personal data is handled

IV. Data on taxonomy

V. Data on spatial geography (**sites of occurrence**)

VI. Data on soil-biodiversity observations

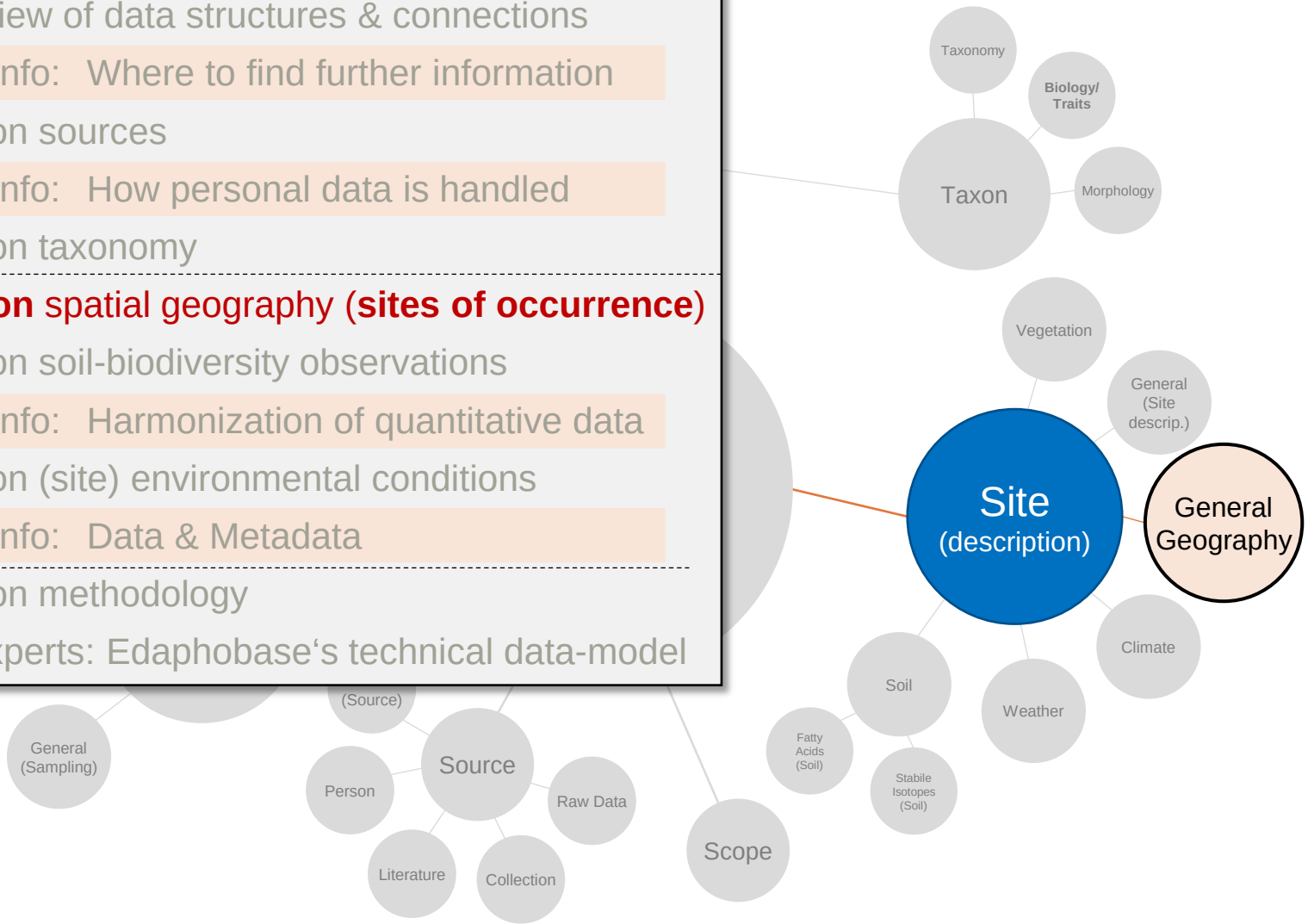
Info: Harmonization of quantitative data

VII. Data on (site) environmental conditions

Info: Data & Metadata

VIII. Data on methodology

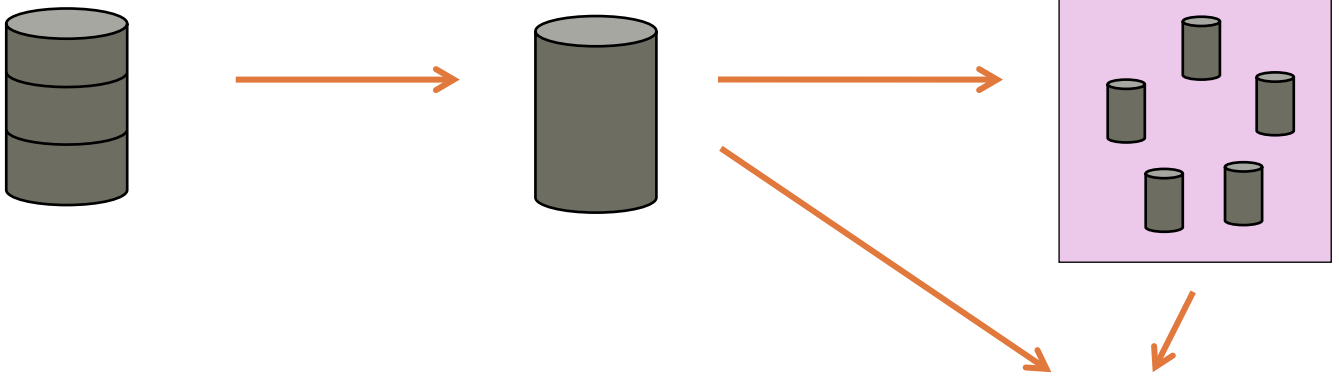
IX. For experts: Edaphobase‘s technical data-model



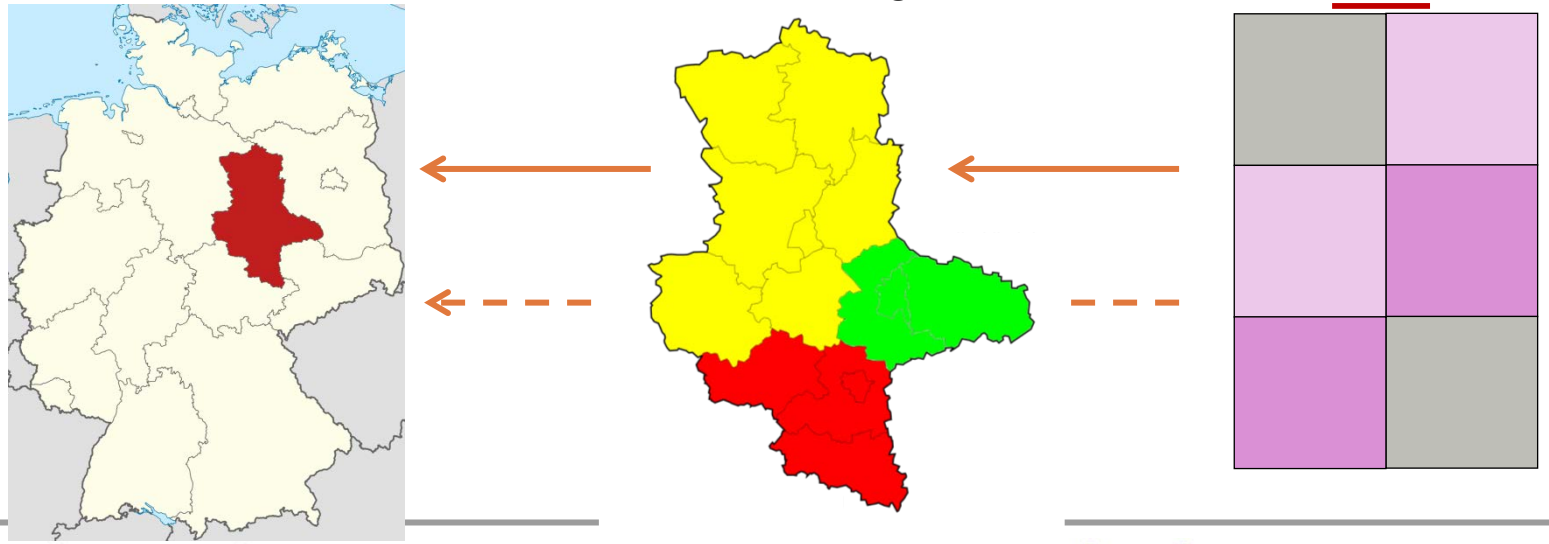
Edaphobase "Geography"

(hierarchal spatial relationships)

- Sub-Sample
- Sample
- Plot



- Country
- „State" / Region
- **Site**



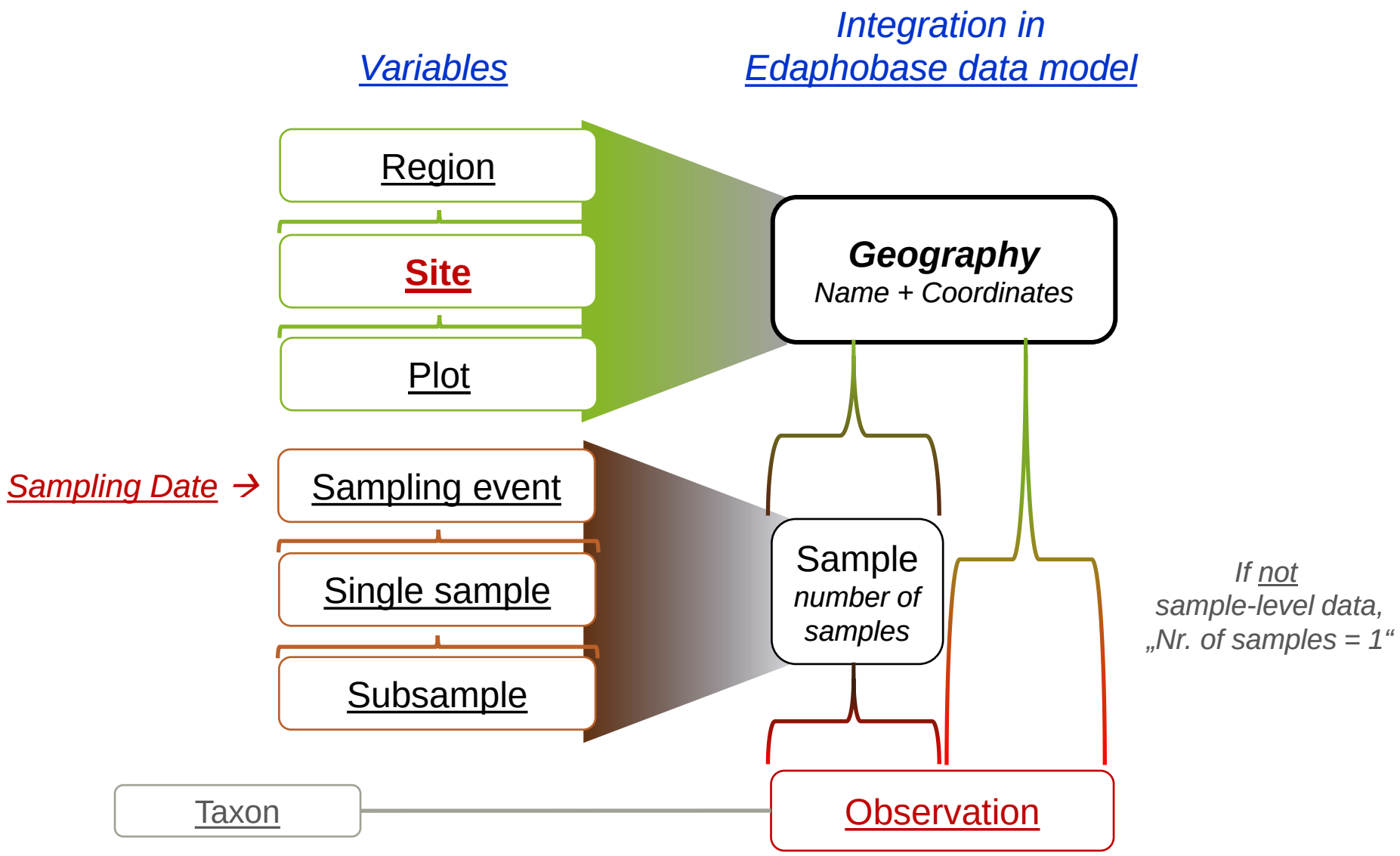
Edaphobase "Geography"

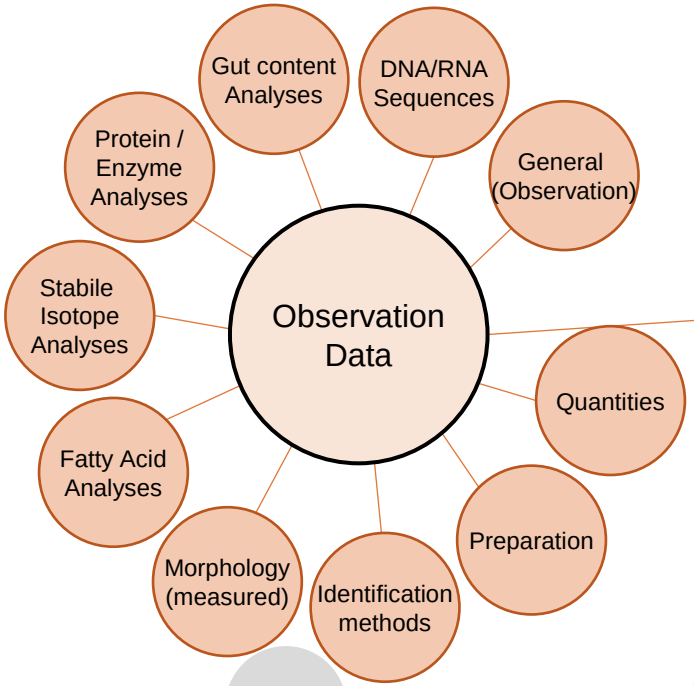
(hierarchal spatial relationships)

- Sub-Sample
- Sample
- Plot
- Coun

- **Geo-coordinates ! (+ "radius" *[precision]*)**
- All data can be added at any hierarchal level
  - coordinates
  - biodiversity
  - site metadata
- **Data harmonization unit = „site"**
  (→ data comparability)

Edaphobase "Geographical hierarchy"





I. Edaphobase‘s „data philosophy“

II. Overview of data structures & connections

Info: Where to find further information

III. Data on sources

Info: How personal data is handled

IV. Data on taxonomy

V. Data on spatial geography (sites of occurrence)

VI. Data on soil-biodiversity observations

Info: Harmonization of quantitative data

VII. Data on (site) environmental conditions

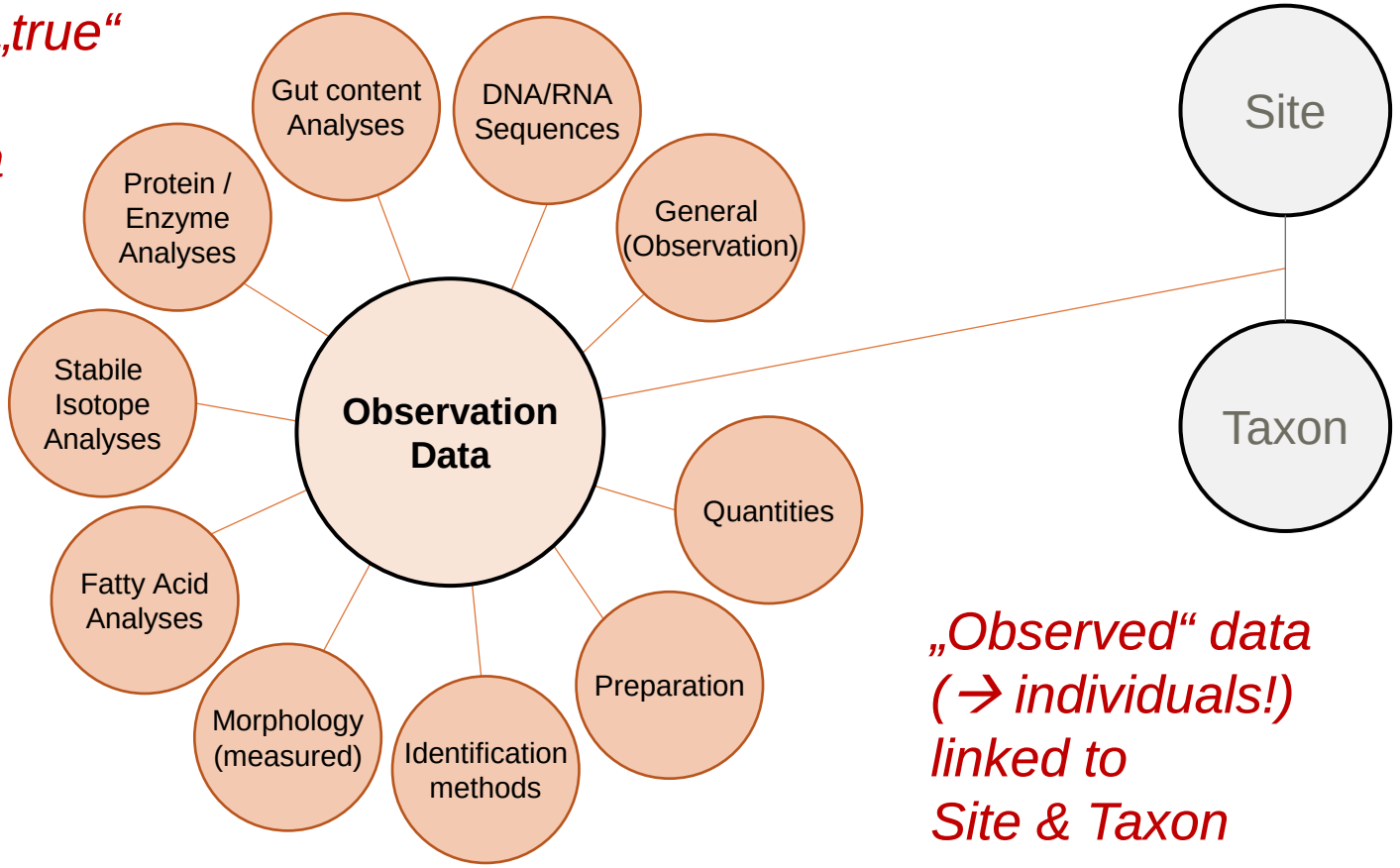
Info: Data & Metadata

VIII. Data on methodology

IX. For experts: Edaphobase‘s technical data-model

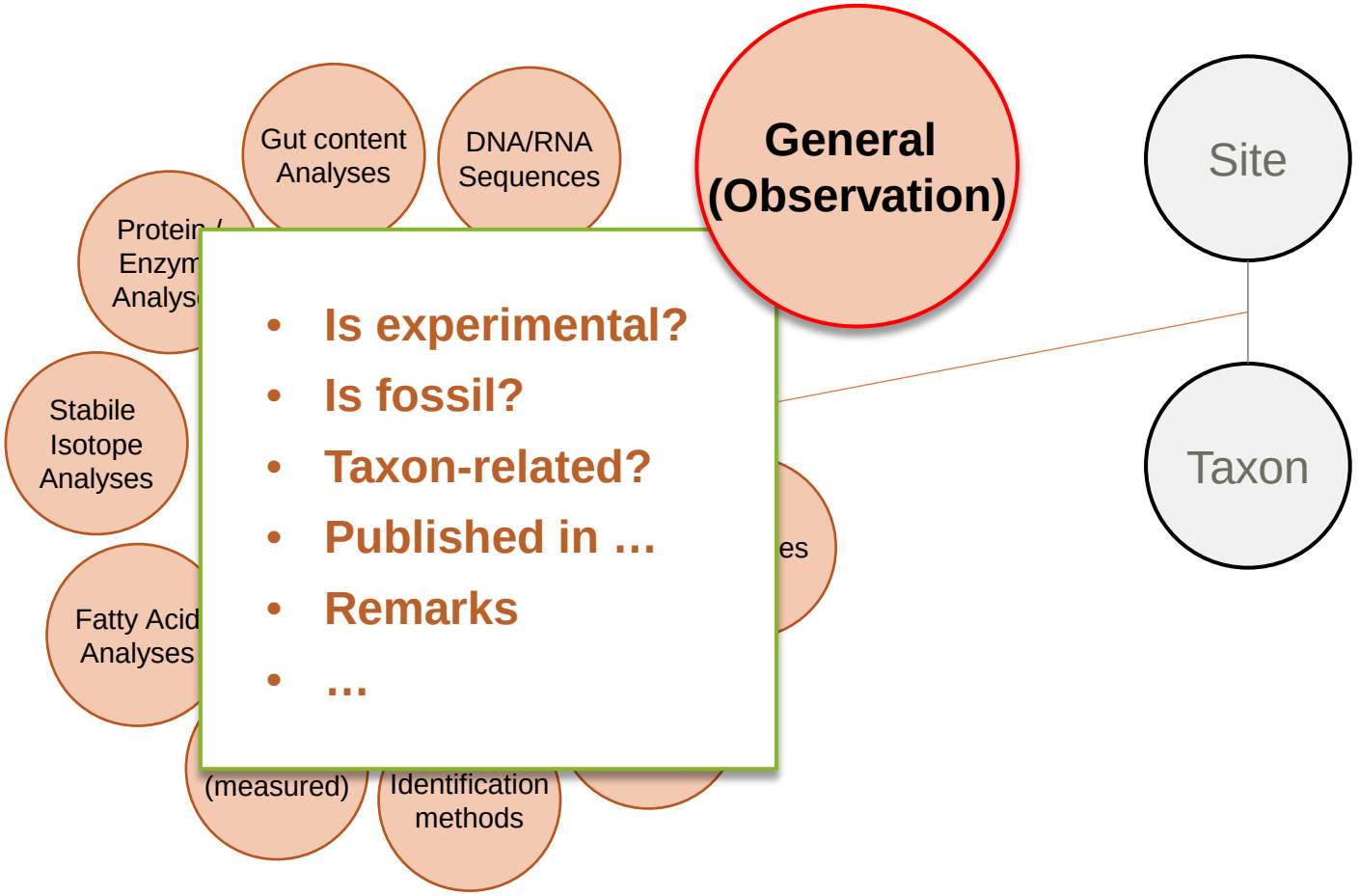
Edaphobase Object / Observation

Represents the „true“ (= sampled) biodiversity data



„Observed“ data (→ individuals!) linked to Site & Taxon

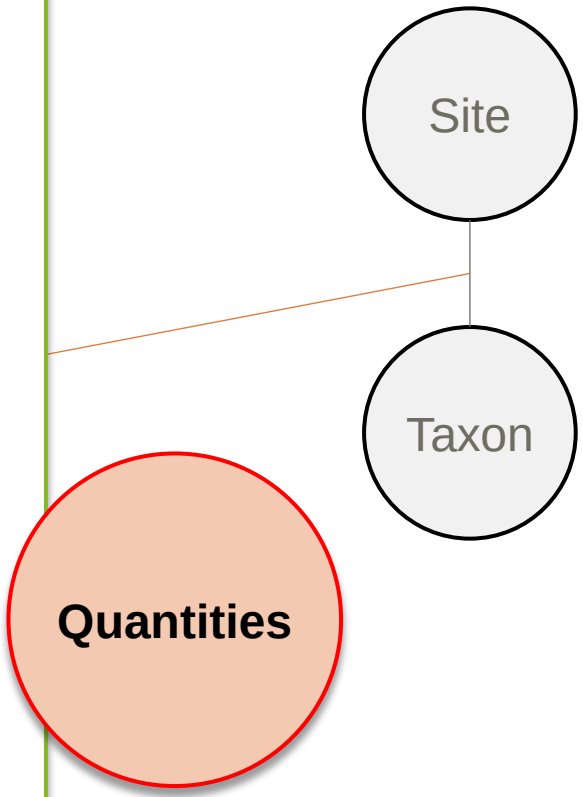
Edaphobase Object / Observation



Edaphobase Object / Observation

This is the important biodiv data!

- **Total ind. in sample *(= Counts)***
- **Count in Collection**
- **Abundance (i.e., Ind./m²)**
- **Activity density (Ind./Trap/Time)**
- **Dominance** or **Frequency**
- **Dry** or **Fresh Weight *(specimen)***
- **Biomass *(total ind./taxon)***
- **Sex**
- **Developmental stage**
- **…**



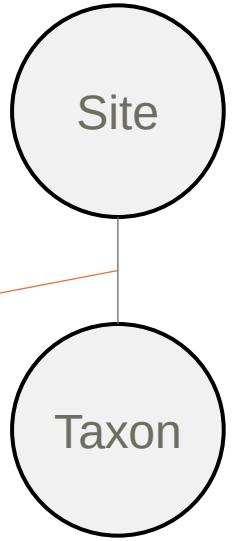
Edaphobase Object / Observation

This is the important biodiv data!

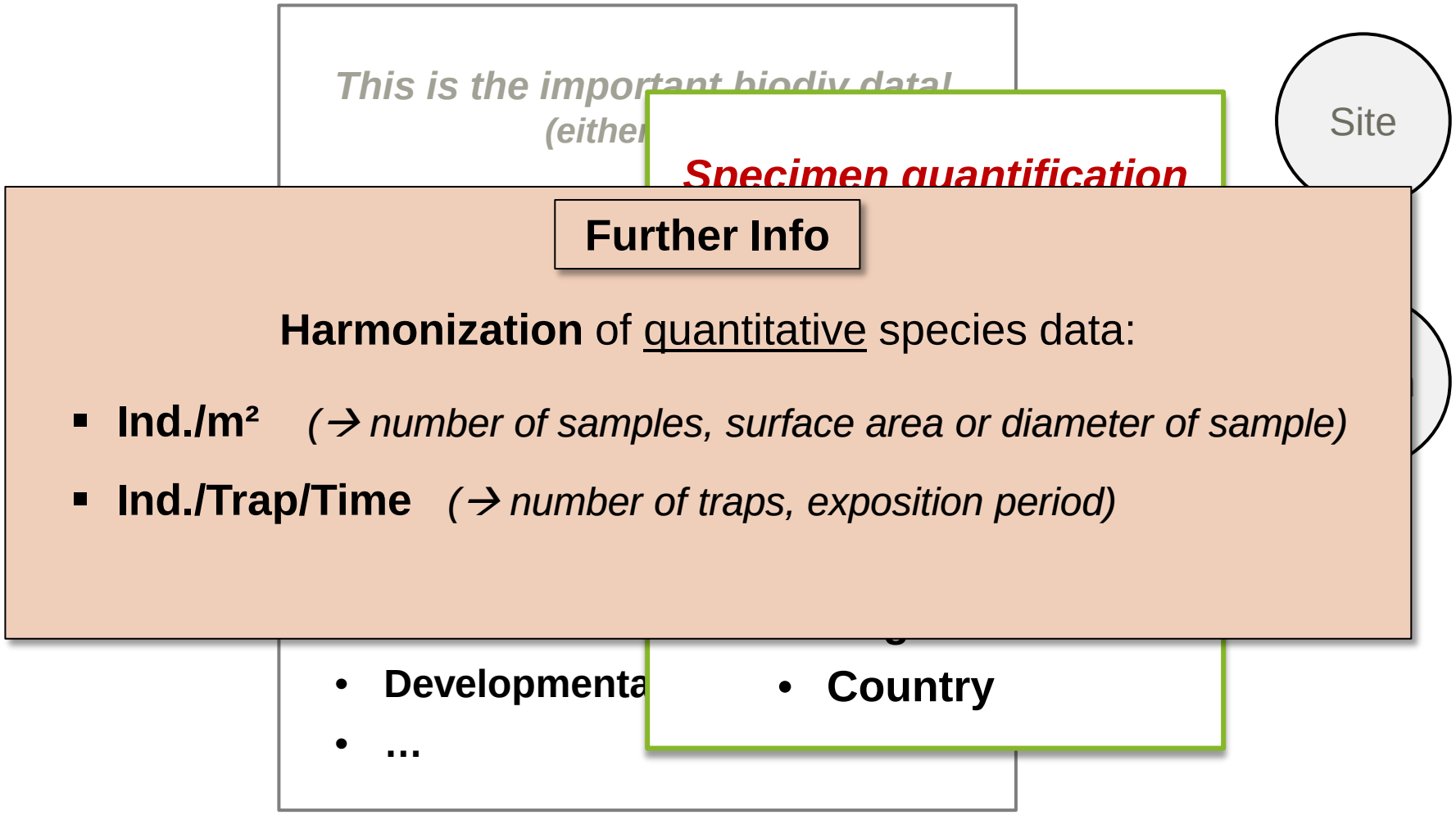
- **Total ind. in sa**
- **Count in Colle**
- **Abundance (i.**
- **Activity densit**
- **Dominance** or
- **Dry** or **Fresh W**
- **Biomass** ***(tota***
- **Sex**
- **Developmenta**
- **…**

Specimen quantification can be linked to <u>any</u> spatial scale !

- **Sub-sample**
- **Sample**
- **Plot**
- **<u>Site</u>**
- **Region**
- **Country**

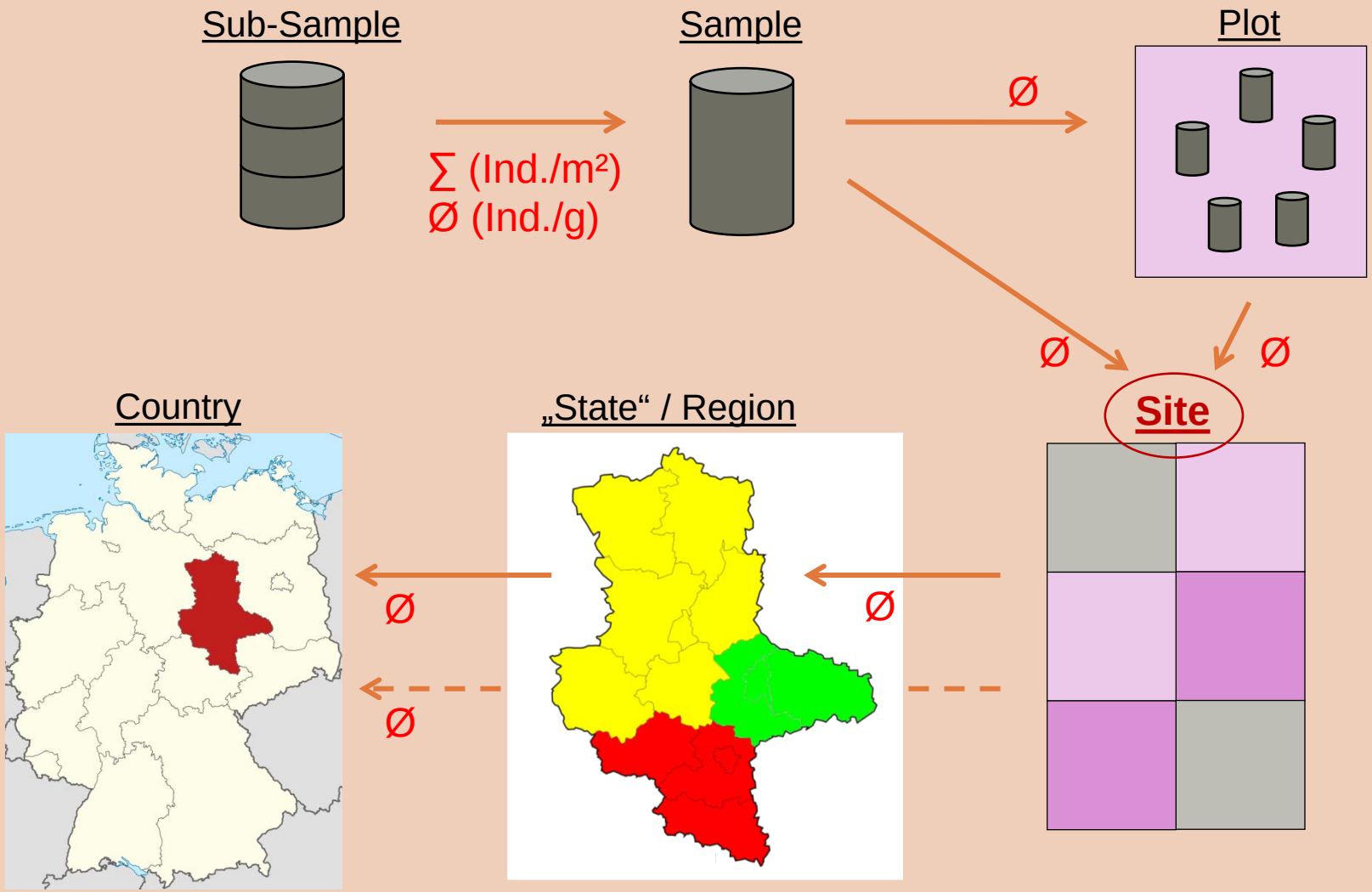


Edaphobase Object / Observation

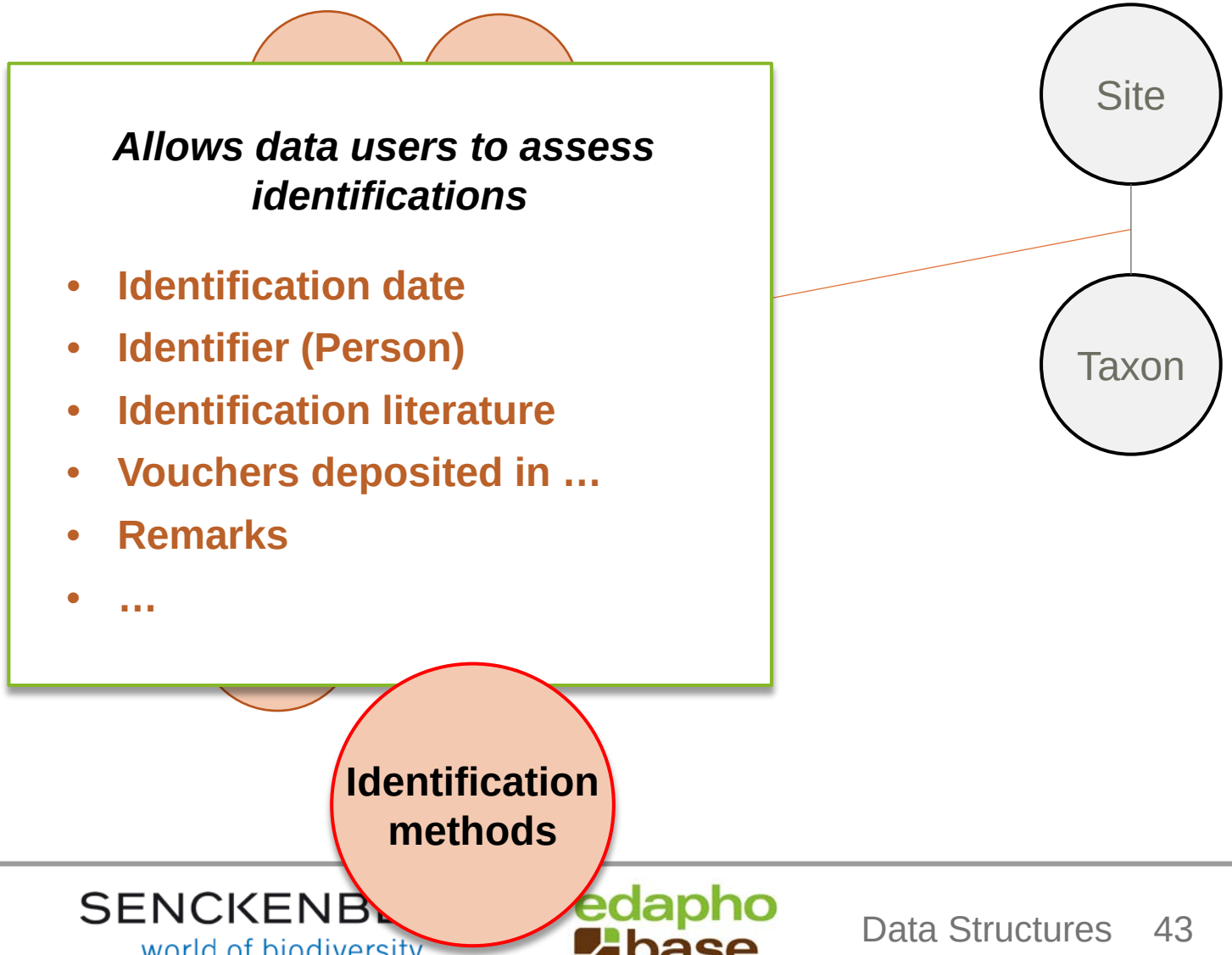


Further Info

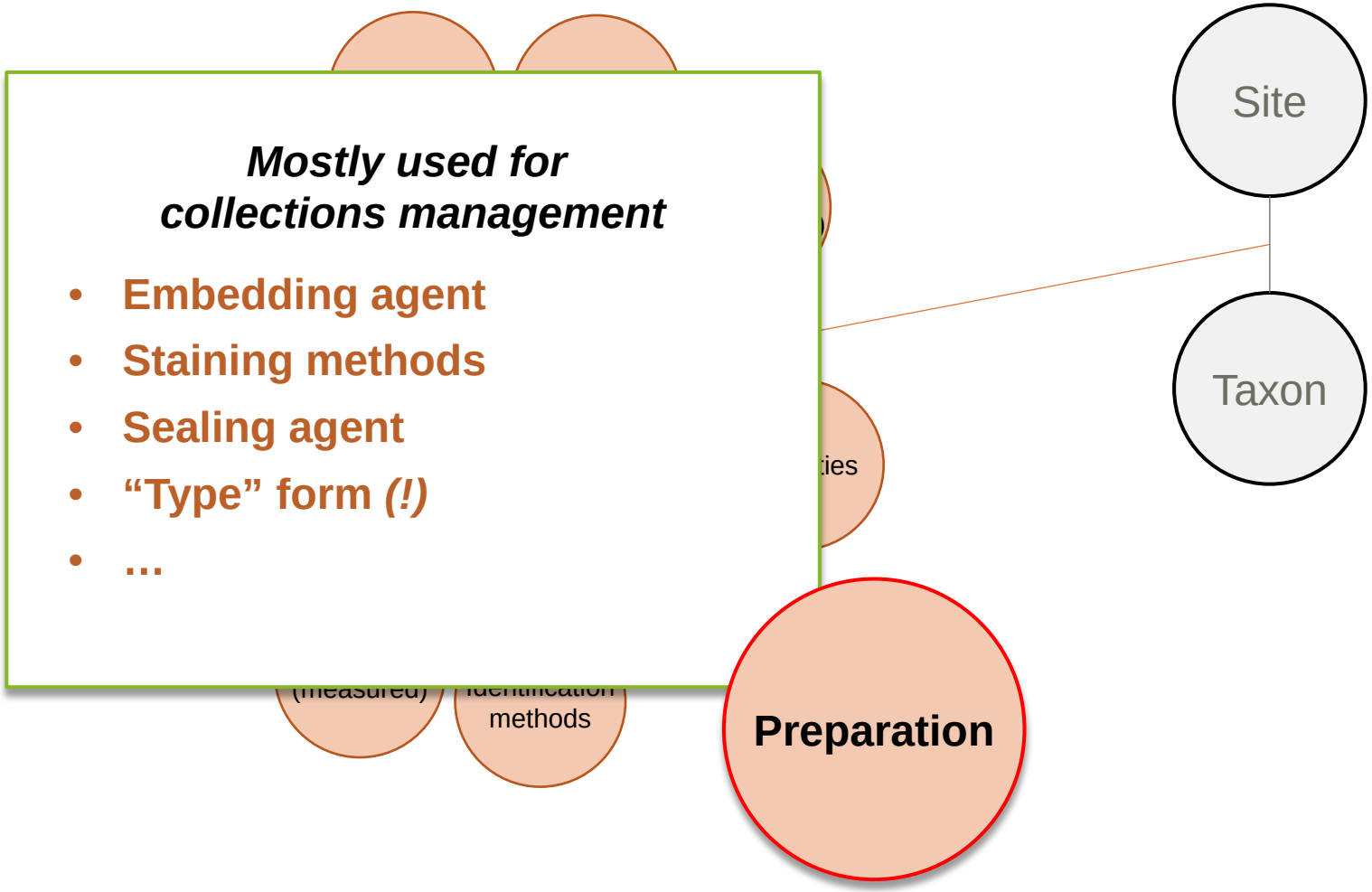
Spatial harmonization of biodiversity quantities



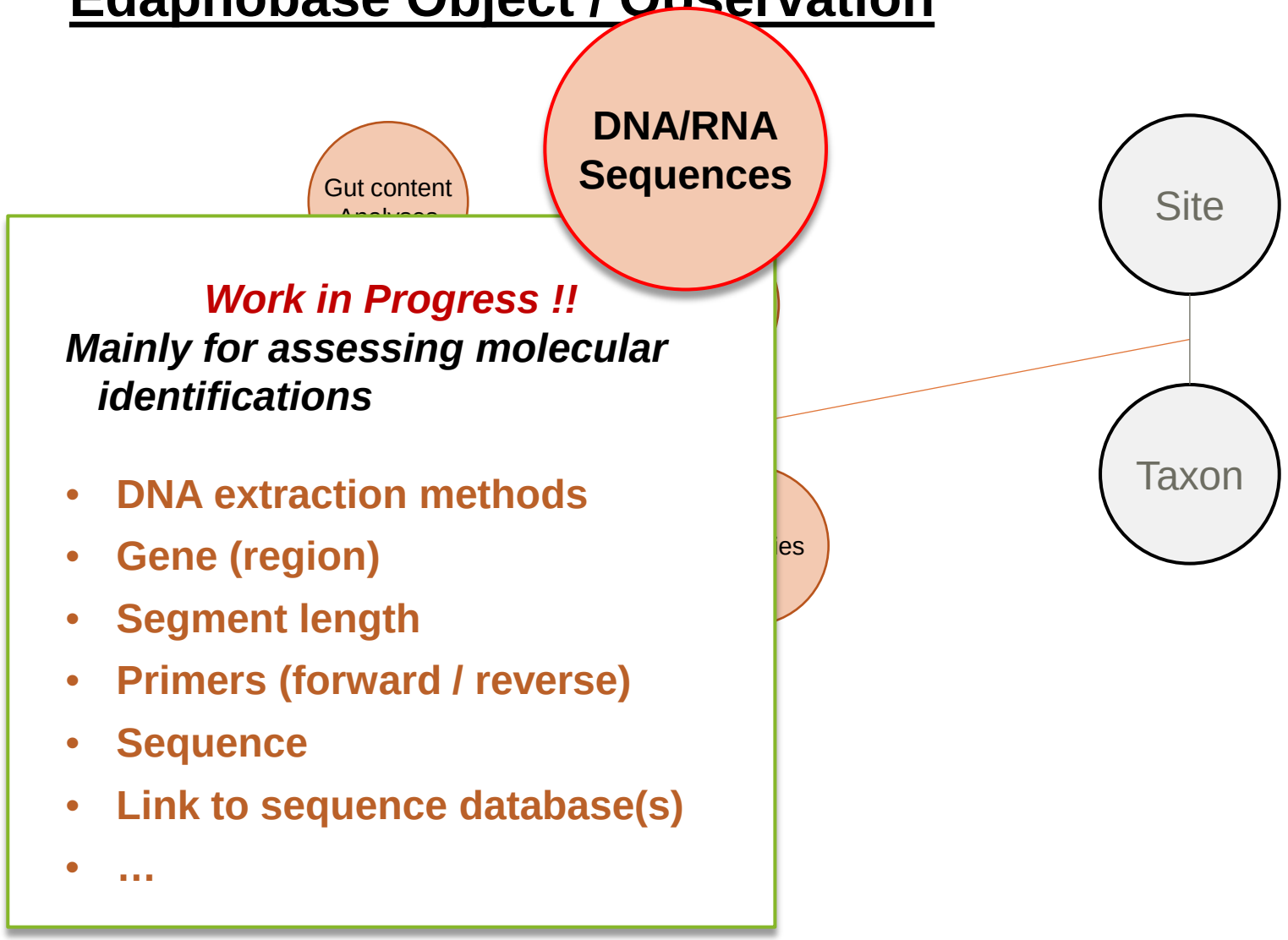
Edaphobase Object / Observation



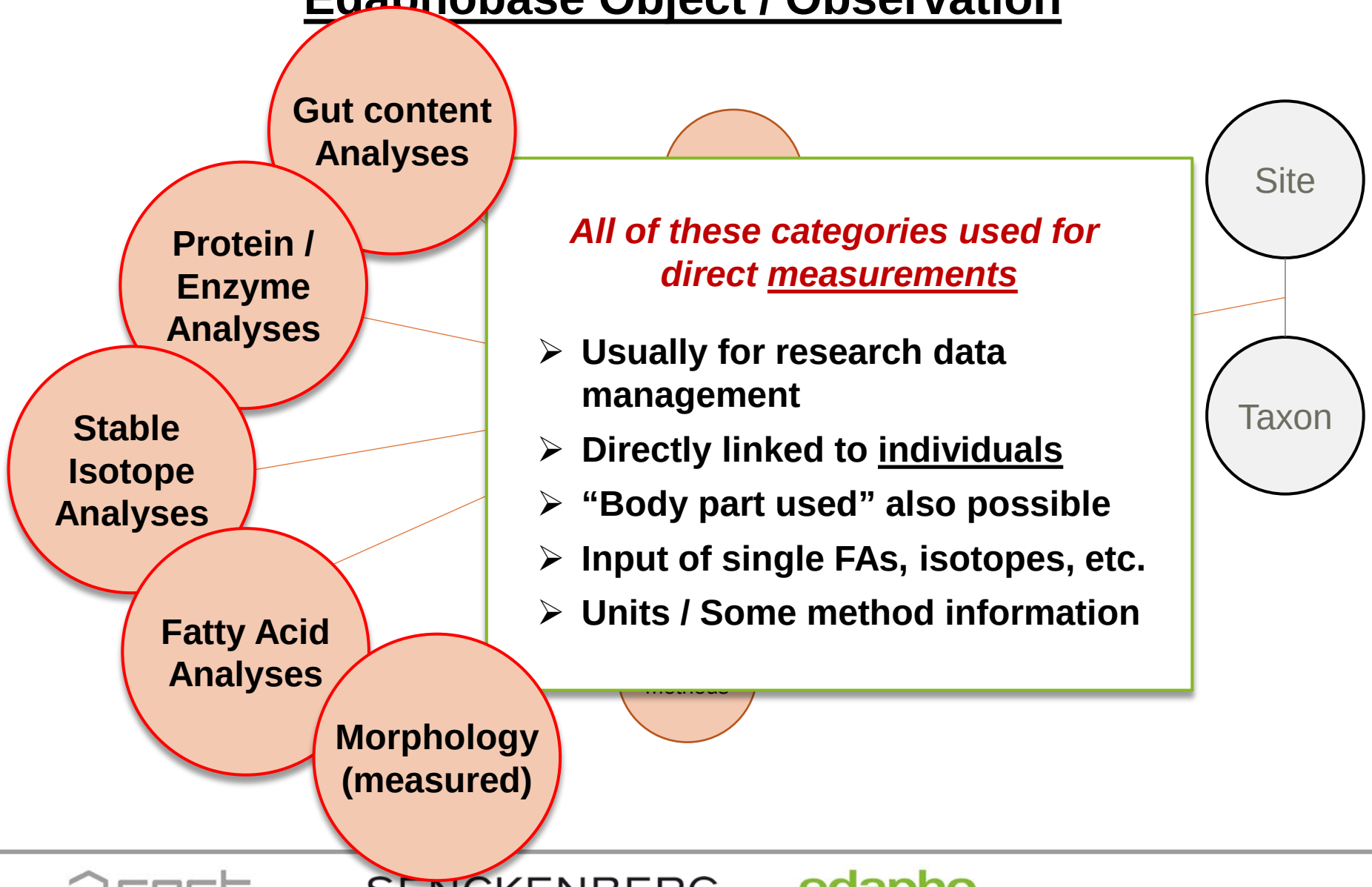
Edaphobase Object / Observation



Edaphobase Object / Observation



Edaphobase Object / Observation



I. Edaphobase‘s „data philosophy“

II. Overview of data structures & connections

Info: Where to find further information

III. Data on sources

Info: How personal data is handled

IV. Data on taxonomy

V. Data on spatial geography (sites of occurrence)

VI. Data on soil-biodiversity observations

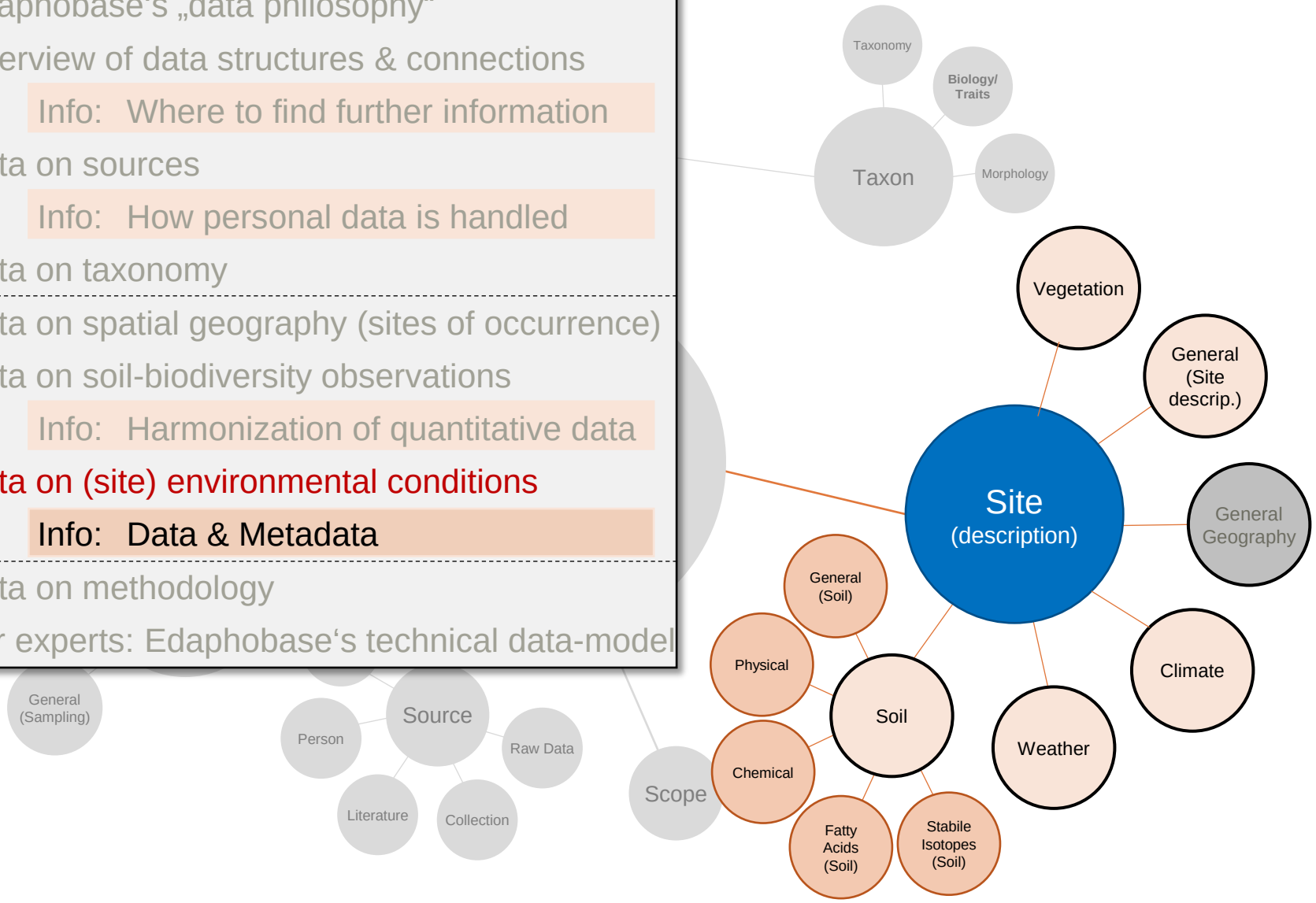
Info: Harmonization of quantitative data

VII. Data on (site) environmental conditions

Info: Data & Metadata

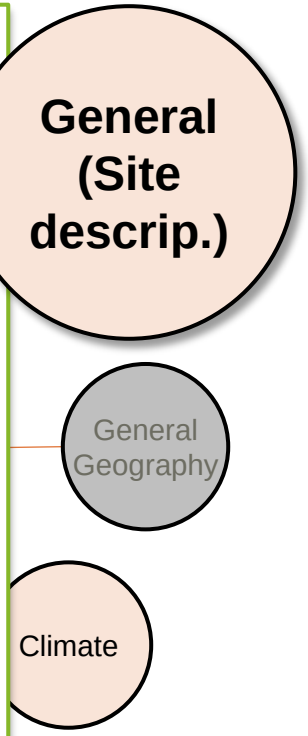
VIII. Data on methodology

IX. For experts: Edaphobase‘s technical data-model



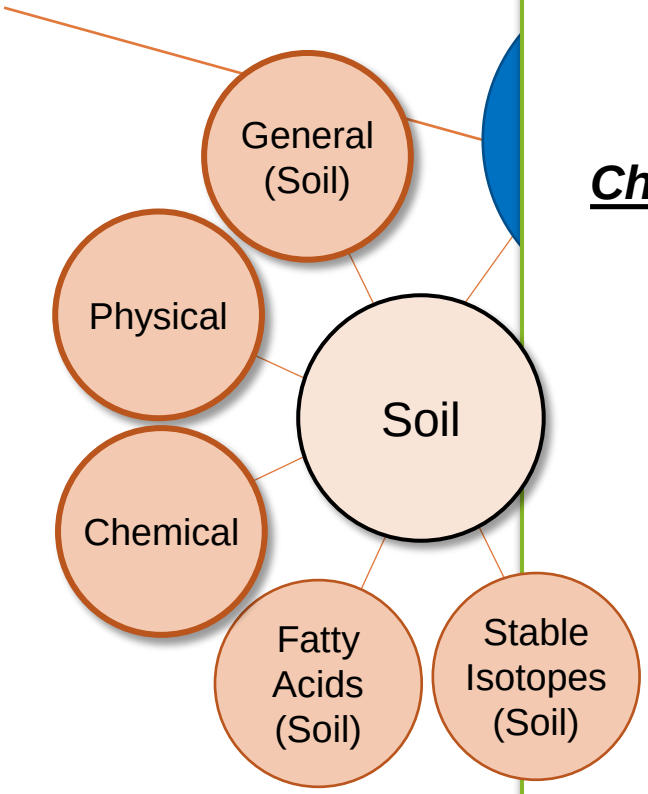
Edaphobase (environmental) Site description

- Description date
- Habitat Type / Land Use
  - *German BfN*
  - *CORINE* [1]
  - *EUNIS* [2]
- Microhabitat
- Altitude (m.a.s.l.)
- Slope
- Monitoring site? (Y/N)
- Is experimental? (Y/N)
- (Anthropogenic) influence ***[WiP]***
- …



[1] https://land.copernicus.eu/pan-european/corine-land-cover/clc2018
[2] https://eunis.eea.europa.eu/

Edaphobase (environmental) Site description



General

- Description date
- Soil Type
- Humus Form
- Sampled Horizon
- Soil depth
- H_2O / WHC
- …

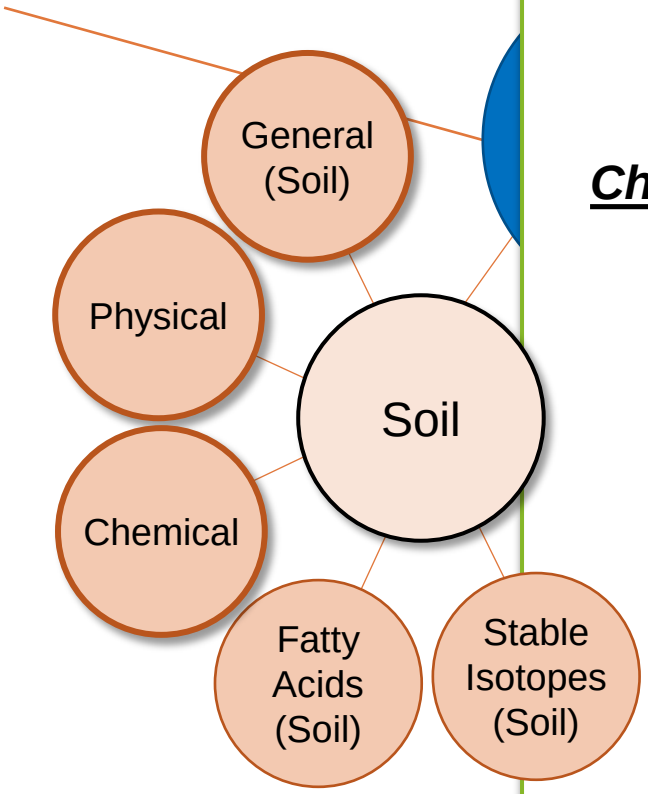
Chemical parameters

- SOM
- $C_{org\ /\ total}$
- $N_{org\ /\ total}$
- C/N
- P_{total}
- CEC / Base Saturation
- $CaCO_3$
- …

Physical parameters

- Soil Texture
- % Sand / Clay / Silt
- Density
- Pore Volume
- …

Edaphobase (environmental) Site description



General

- Description date
- Soil Type
- Humus Form
- Sampled Horizon
- Soil depth
- H_2O / WHC

Includes Methods!

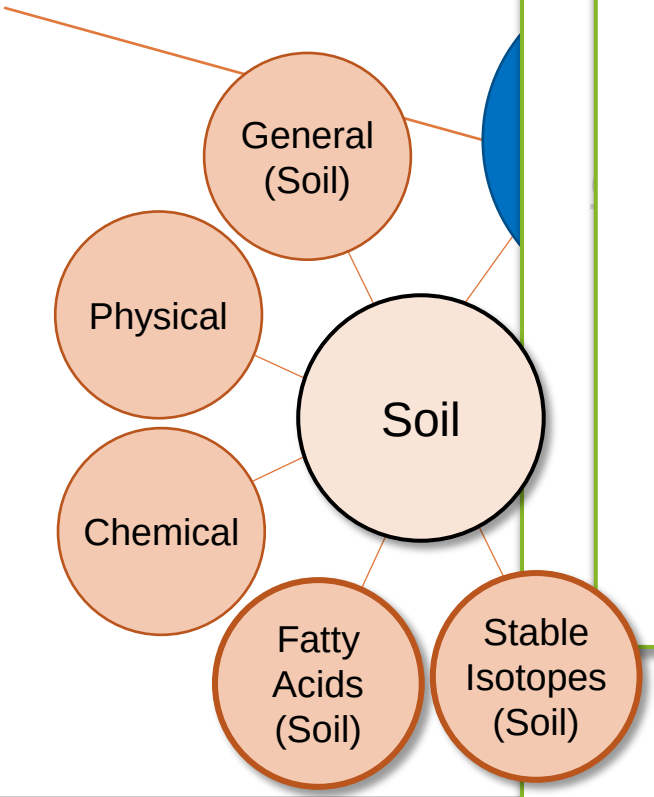
Chemical p… …l parameters

- SOM
- $C_{org / total}$
- $N_{org / total}$
- C/N
- P_{total}
- CEC / Base Saturation
- $CaCO_3$
- …

…oil Texture

- % Sand / Clay / Silt
- Density
- Pore Volume
- …

Edaphobase (environmental) Site description



General

Fatty Acids & Stabile Isotopes

→ **Soil !! (for comparison with animals)**
→ **Generally for RDM**

- **Date**
- **Person**
- **Substrate**
- **Methods**
- **δ15N / δ13C Baselines**
- **Total PFLA**
- **Bacteria : Fungi ratio**
- **Individual FAs**

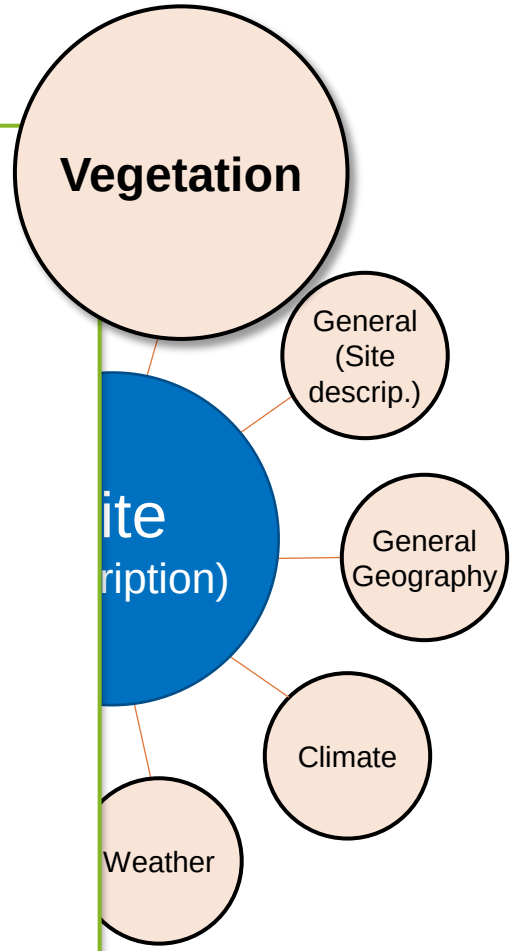
Saturation

– $CaCO_3$

– **…**

Edaphobase (environmental) Site description

- Description date
- Person
- Phytosociological unit
- Total Cover [%]
- Species / Cover [%]
  - Canopy
  - Shrub
  - Herbaceous
  - Moss
  - Epiphytes
- …

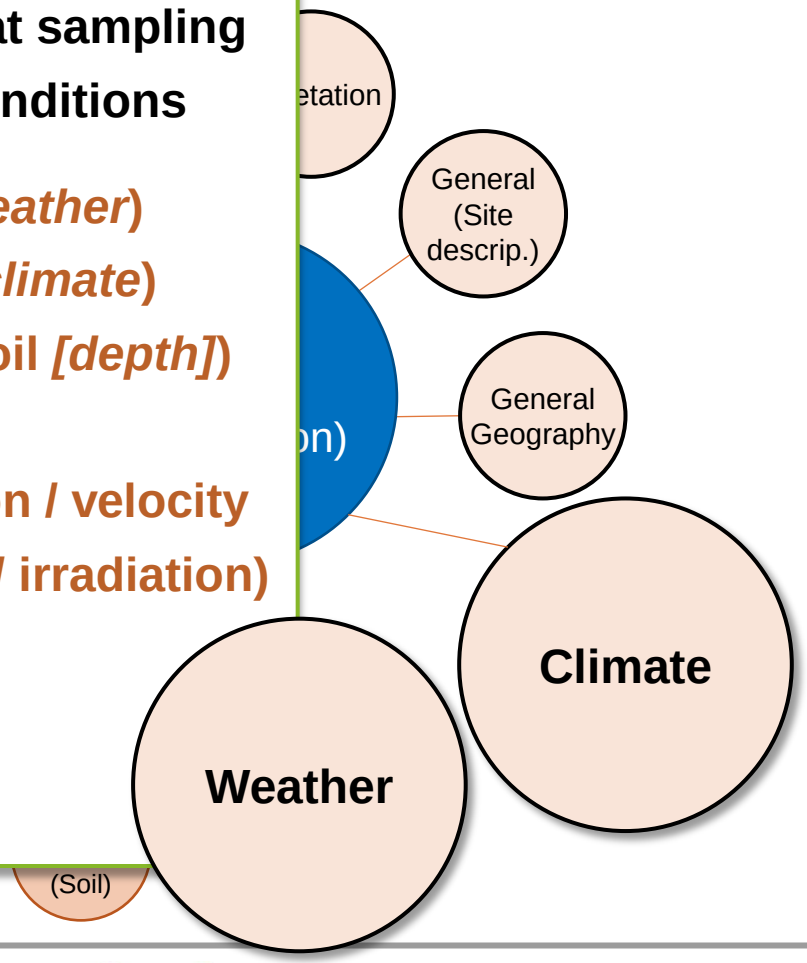


Edaphobase (environmental) Site description

Weather = conditions at sampling

Climate = long-term conditions

- Description date (*weather*)
- Measurement site (*climate*)
- Temperature (air / soil *[depth]*)
- Precipitation
- (main) Wind direction / velocity
- Sunshine (duration / irradiation)
- Frost days
- …



Further Info

„Data“ & „Metadata“

„*philosophical*“ → one person‘s metadata is another‘s data

„*metadata*“: one global value for the entire dataset
„*data*“: different values *(in a column)* for different data records

Edaphobase can accommodate both viewpoints !

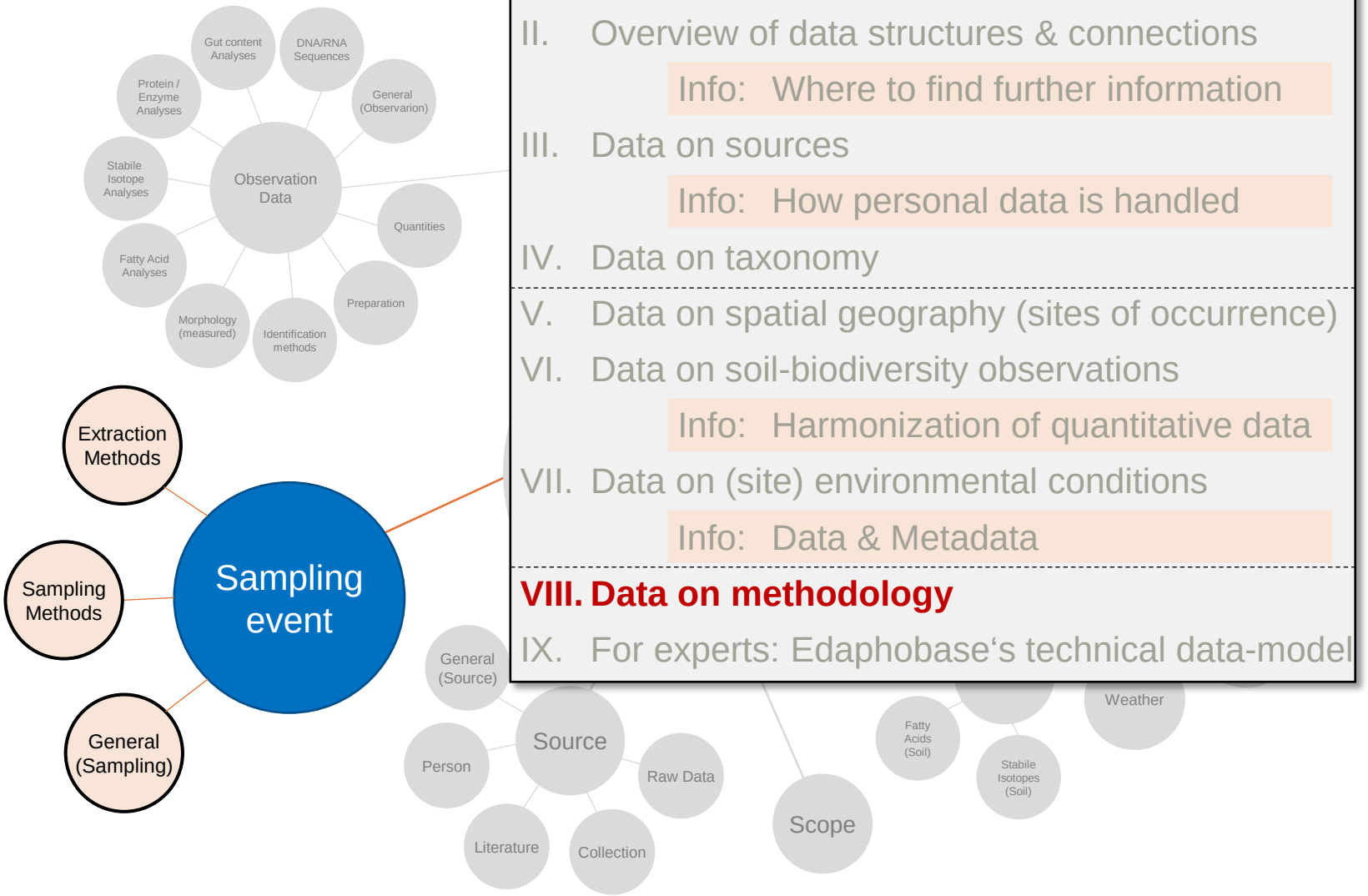
- (almost) every variable can be either “data” or “metadata”

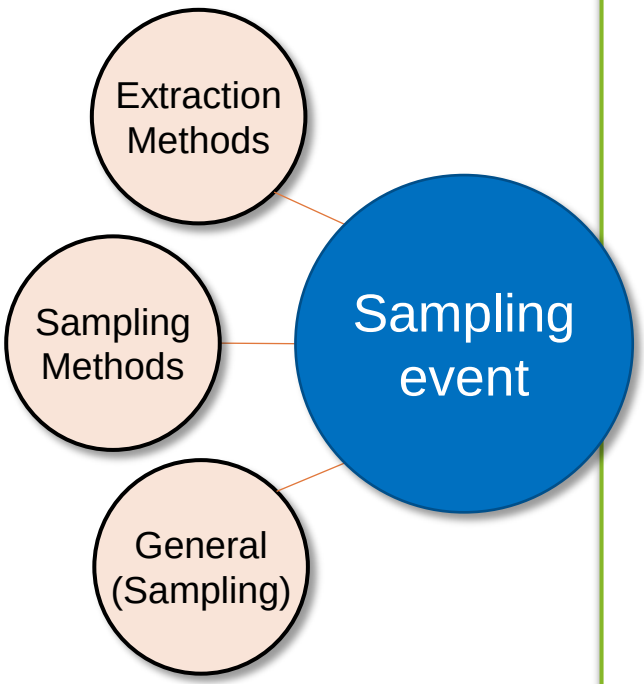
Regarding Metadata

Edaphobase follows, i.e., Inspire, DataCite, BonaRes, etc.

I. Dataset Metadata → describes the dataset (title, owner, etc.)
II. Environmental Metadata → describes the sites of occurrence
III. Methodological Metadata → describes the sampling methods

→ Tab „*Readme – Metadata*“ in *Edaphobase-Information fields_2019v_EN_public version.xlsx*


Gut content Analyses
DNA/RNA Sequences
Protein / Enzyme Analyses
General (Observarion)
Stabile Isotope Analyses
Observation Data
Quantities
Fatty Acid Analyses
Preparation
Morphology (measured)
Identification methods
Extraction Methods
Sampling Methods
Sampling event
General (Sampling)
General (Source)
Source
Person
Raw Data
Literature
Collection
Scope
Fatty Acids (Soil)
Stabile Isotopes (Soil)
Weather
I. Edaphobase‘s „data philosophy“
II. Overview of data structures & connections
Info: Where to find further information
III. Data on sources
Info: How personal data is handled
IV. Data on taxonomy
V. Data on spatial geography (sites of occurrence)
VI. Data on soil-biodiversity observations
Info: Harmonization of quantitative data
VII. Data on (site) environmental conditions
Info: Data & Metadata
VIII. Data on methodology
IX. For experts: Edaphobase‘s technical data-model



General

- Observation date (sampling event)
- Sample code
- Remarks
- …

Sampling methods

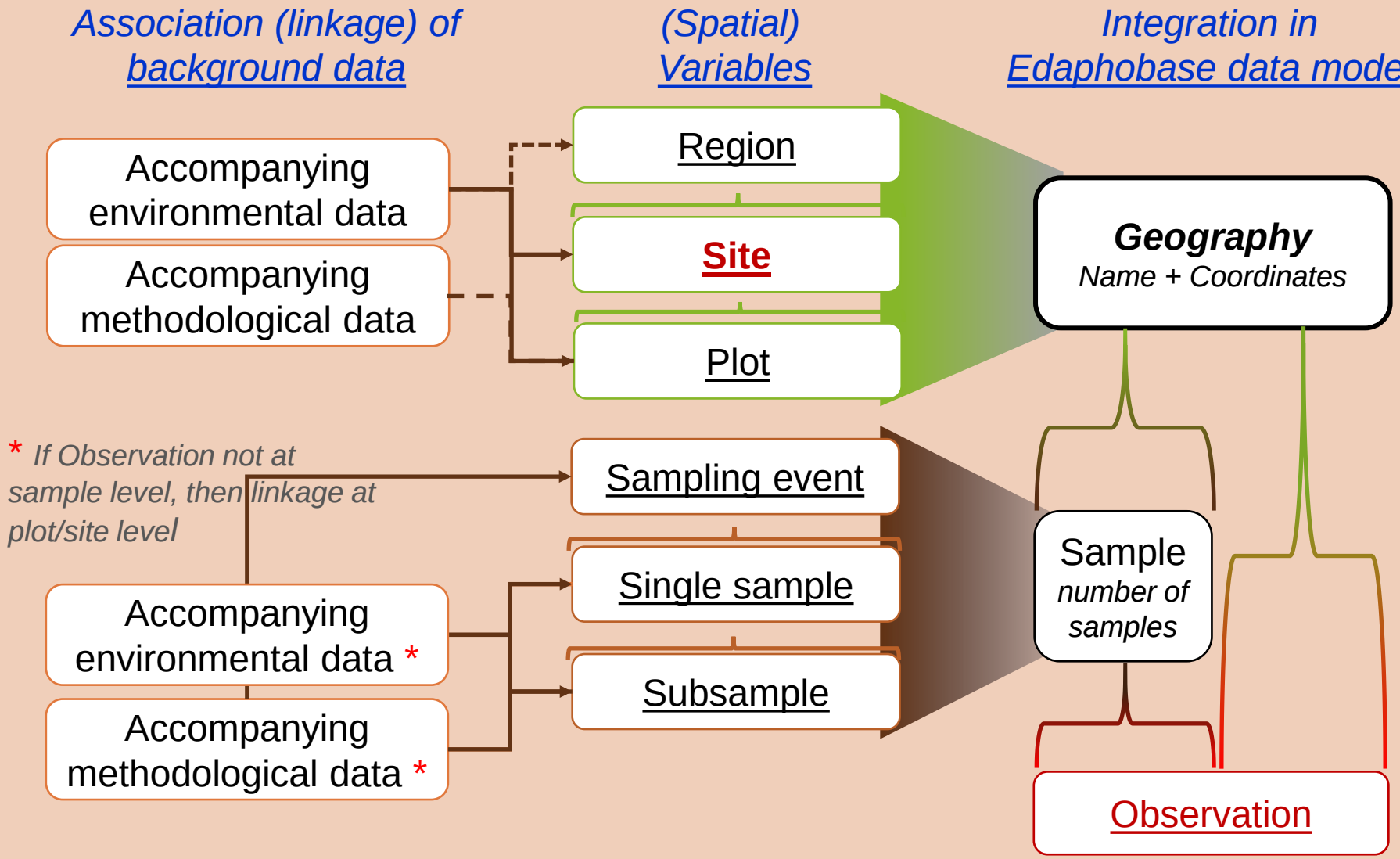
- Sampling method (→ *from list*)
- Standardization (ISO, DIN, etc.)
- Number of samples
- Sample size, area / Exposition time
- Sampled horizon / depth
- …

Extraction methods

- Extraction method (→ *from list*)
- Duration / Temp. gradient
- Conservation / Fixation medium
- …

Further Info

Linking environmental & methodological (meta)data to sites



In conclusion

Edaphobase maintains *very* many data fields (> 600 !)

= Options !! *(not all mandatory)*

→ *What is actually required?*!

Minimum data set: *(i.e., useful for determining species distributions)*

(→ "what", "where", "when"?)

Species, sites (incl. geo-coordinates!), sampling dates

Minimum = presence/absence data

Recommended data set: (i.e., can be used for ecological analyses)

(→ + "under what environm. conditions", "how", "by whom"?)

minimum data set +

habitat types / land-use, soil parameters

sampling methods, who sampled and determined the taxa

Recommended = abundance data

SENCKENBERG
world of biodiversity

edapho base

I. Edaphobase‘s „data philosophy“
II. Overview of data structures & connections
Info: Where to find further information
III. Data on sources
Info: How personal data is handled
IV. Data on taxonomy

V. Data on spatial geography (sites of occurrence)
VI. Data on soil-biodiversity observations
Info: Harmonization of quantitative data
VII. Data on (site) environmental conditions
Info: Data & Metadata

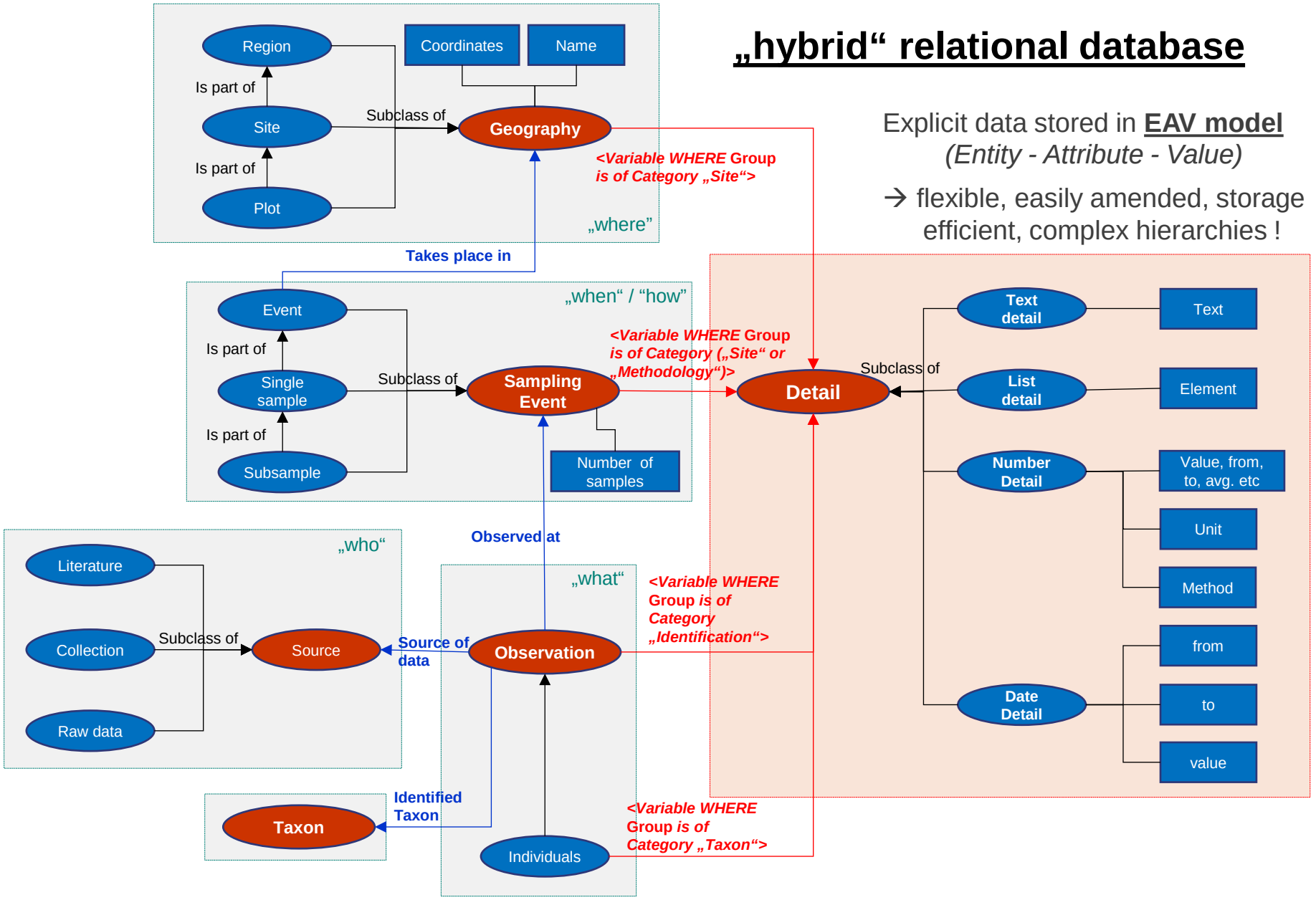
VIII. Data on methodology
IX. For experts: Edaphobase‘s technical data-model

cost
EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

„hybrid“ relational database

Explicit data stored in **EAV model**
(Entity - Attribute - Value)

→ flexible, easily amended, storage efficient, complex hierarchies !



→ EAV Data Model (explanation)

Example: common „flat“ table

Site	Habitat	Sample	pH	C/N	Date
1	Forest	1	4.9-5.2		2018-01-01
1	Forest	2	7.0	10	2018-03-21
1	Forest	3			2018-04-24 2018-06-07

Example: standard EAV table

Entity	Attribute	Value
Sample 1	pH	"4.9-5.2"
Sample 1	Date	"2018-01-01"
Sample 2	pH	"7.0"
Sample 2	C/N	"10"
Sample 2	Date	"2018-03-21"
Sample 3	Date	"2018-04-24"
Sample 3	Date	"2018-06-07"
Site	Habitat	"Forest"

→ Edaphobase‘s EAV/CR (*with Class and Relationships*) table

ID	Entity	Attribute	Value *exact*	Value *min*	Value *max*	Value *date*	Value *list entry*
10511	Sample 1	pH		4.9	5.2		
10512	Sample 1	Date				2018-01-01	
10521	Sample 2	pH	7				
10523	Sample 2	C/N	10				
10522	Sample 2	Date				2018-03-21	
10532	Sample 3	Date				2018-04-24	
105322	Sample 3	Date				2018-06-07	
10501	Site 1	Habitat					Forest

Edaphobase is programmed in:

- Database itself:
  PostgreSQL 12
- Query Portal:
  PHP, Javascript, PostGIS 2.2., Apache HTTP Server 2.4
- Data Upload Software (Wizard):
  Java 8 (written in Eclipse iDE)

The software code is open source and freeware
(but the database software is © Senckenberg Society for Nature Research)

Growing ideas through networks

Cost Action CA 18237

EUdaphobase

Upcoming online training courses:

- **Use of Edaphobase (Query Portal)**
- **Use of Data-Upload “Wizard”**

David Russell, Stephan Lesch, Sebastian Rick (Senckenberg, Germany)

cost EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

Funded by the Horizon 2020 Framework Programme of the European Union

SENCKENBERG world of biodiversity

Cost Action CA 18237

EUdaphobase

ed[medien]

and especially ed[medien] Görlitz (https://edmedien.de) for the friendly, technical support !

Edaphobase Data Structures

Online Training Course (”Webinar”®)

David Russell, Stephan Lesch, Sebastian Rick (Senckenberg, Germany)

cost EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

Funded by the Horizon 2020 Framework Programme of the European Union

SENCKENBERG world of biodiversity