

ITCER Workshop, Oct 27th to 28th 2022

Expert Workshop on GBIF Data



Agenda

October, 28th 2022

- 10.00 SNSB as GBIF data publisher with DWB-BioCASE data publication pipelines
- 10.30 National Museums of Kenya as GBIF data publisher
- 11.00 Discussion board: GBIF distributed infrastructure for publishing biodiversity data: IPT, Darwin Core, ABCD, technical interoperability
- 11.30 ITCER research, teaching and data
- 12.30 Lunch break
- 14.00 Discussion board: Regional data capture and data mobilization at ITCER Training Centre
- 14.30 Discussion board (with 3 min introduction): GBIF data model, GBIF dataset classes, data publishing and data papers, datasets from universities, agencies and business sector for ecological and environmental analyses
- 15.00 General discussion (with 3 min introduction): "how to proceed..."
- 15.30 Results of the meeting (minutes)
- 16.00 End of day 2
- 18.30 Dinner



ITCER:

- Biodiversity Site



- Laboratory for Analyses





ITCER:

- Biodiversity Site
- Laboratory for Analyses
- Data Node
- Outreach Center and
Communication Platform



→ **Research and Teaching**

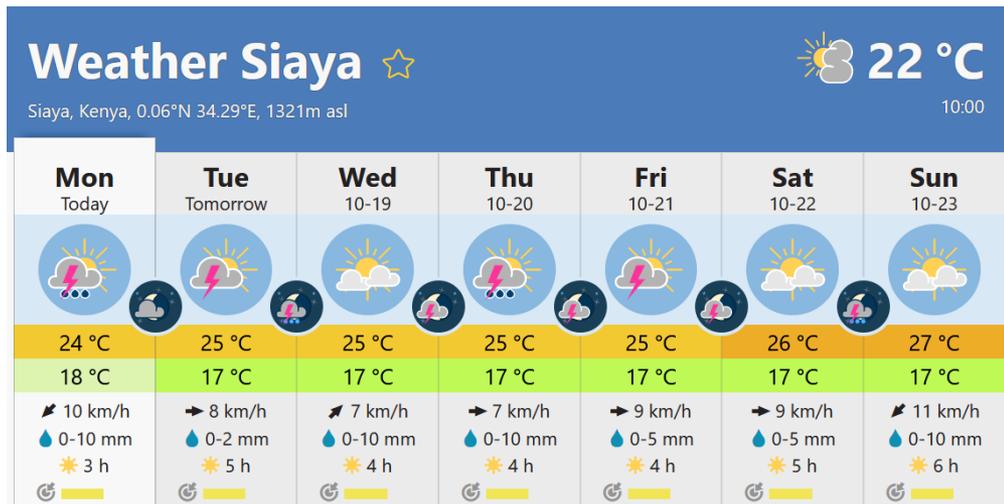
Research and Teaching Activities



- Gathering Contextual Data
- Monitoring
 - Vegetation
 - Insects
 - Mycobiomes
- Grassland Maintenance Experiments

Research and Teaching Activities

- Gathering Contextual Data (Weather Station)



Research and Teaching Activities



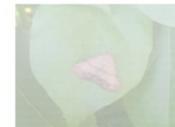
- Monitoring
 - Vegetation
 - Images
 - Logging and streaming
 - Species barcoding
 - Metabarcoding
 - Insects
 - Images
 - Species barcoding
 - Metabarcoding
 - Microbes
 - Images
 - Species barcoding
 - Metabarcoding
 - Secondary metabolite screening

Research and Teaching Activities

- Monitoring on defined vegetation plots
 - Alpha diversity of organisms



DM_DSC09439.JPG_2022-09-24.jp
g



DM_DSC09442.JPG_2022-09-24.jp
g



DM_DSC09465.JPG_2022-09-24.jp
g



DM_DSC09468.JPG_2022-09-24.jp
g



DM_DSC09488.JPG_2022-09-26.jp
g



DM_DSC09489.JPG_2022-09-26.jp
g



DM_DSC09510.JPG_2022-09-26.jp
g



DM_DSC09527.JPG_2022-09-26.jp
g



DM_DSC09537.JPG_2022-09-26.jp
g



DM_DSC09543.JPG_2022-09-26.jp
g



DM_DSC09565.JPG_2022-09-26.jp
g



DM_DSC09567.JPG_2022-09-26.jp
g



DM_DSC09575.JPG_2022-09-26.jp
g



DM_DSC09590.JPG_2022-09-26.jp
g



DM_DSC09597.JPG_2022-09-26.jp
g

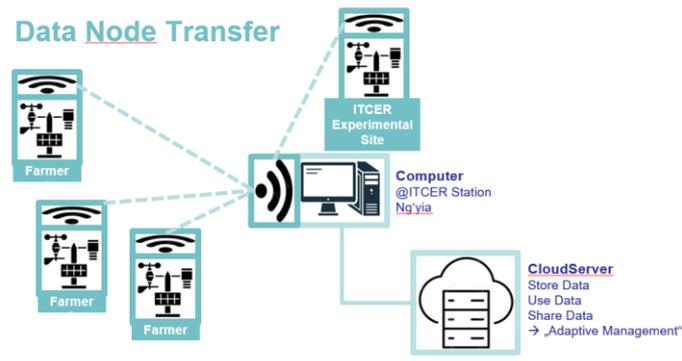
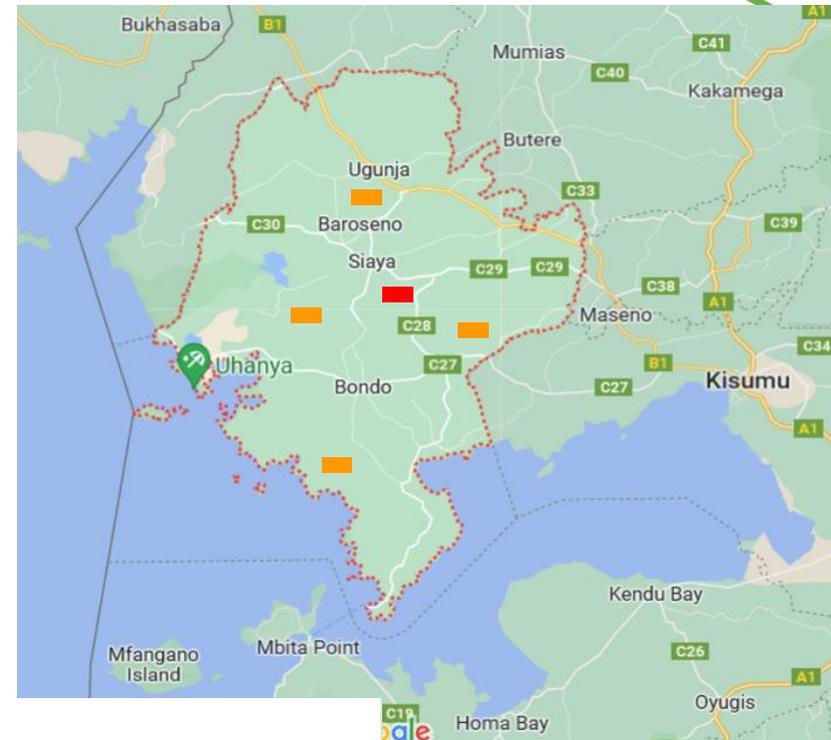
Research and Teaching Activities

- Monitoring of aquatic habitat
 - Alpha diversity of organisms



Research and Teaching Activities

- Location and size
 - Siaya County
 - Siaya township
- Scaling up
 - Comparative studies on grassland biomass production and biodiversity
- Rental
 - Plots of ½ to 1 ha
 - Alternatively: for free from government



Research and Teaching Activities

- Universal Unique Identifier (UUID) and Operation design code (ODC)

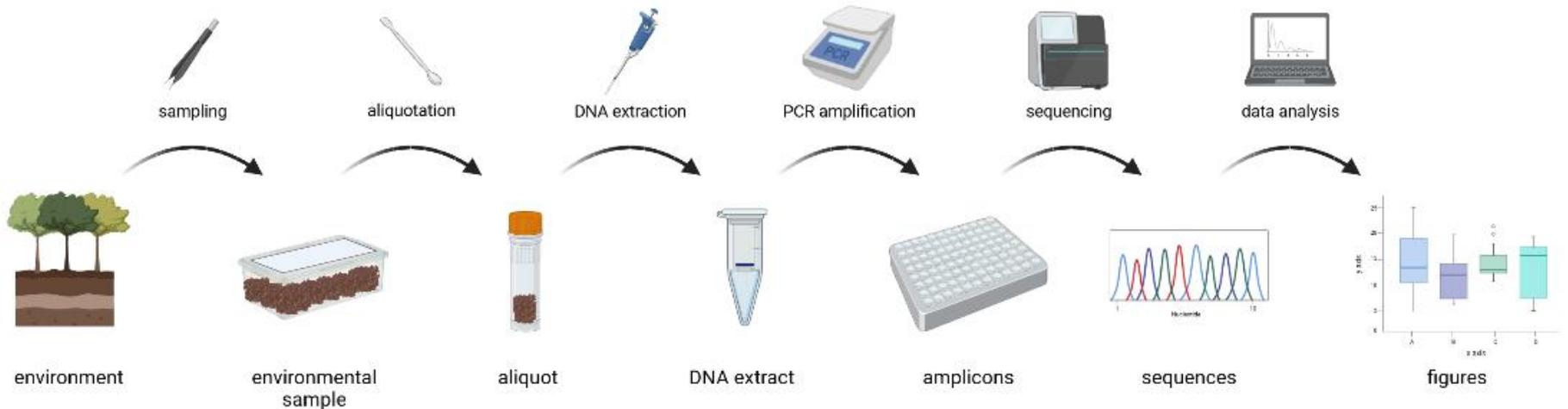
for determination of sample identity and origin

- **Labels** attached to **containers** of environmental **samples**



Research and Teaching Activities

- Metabarcoding Workflow Overview



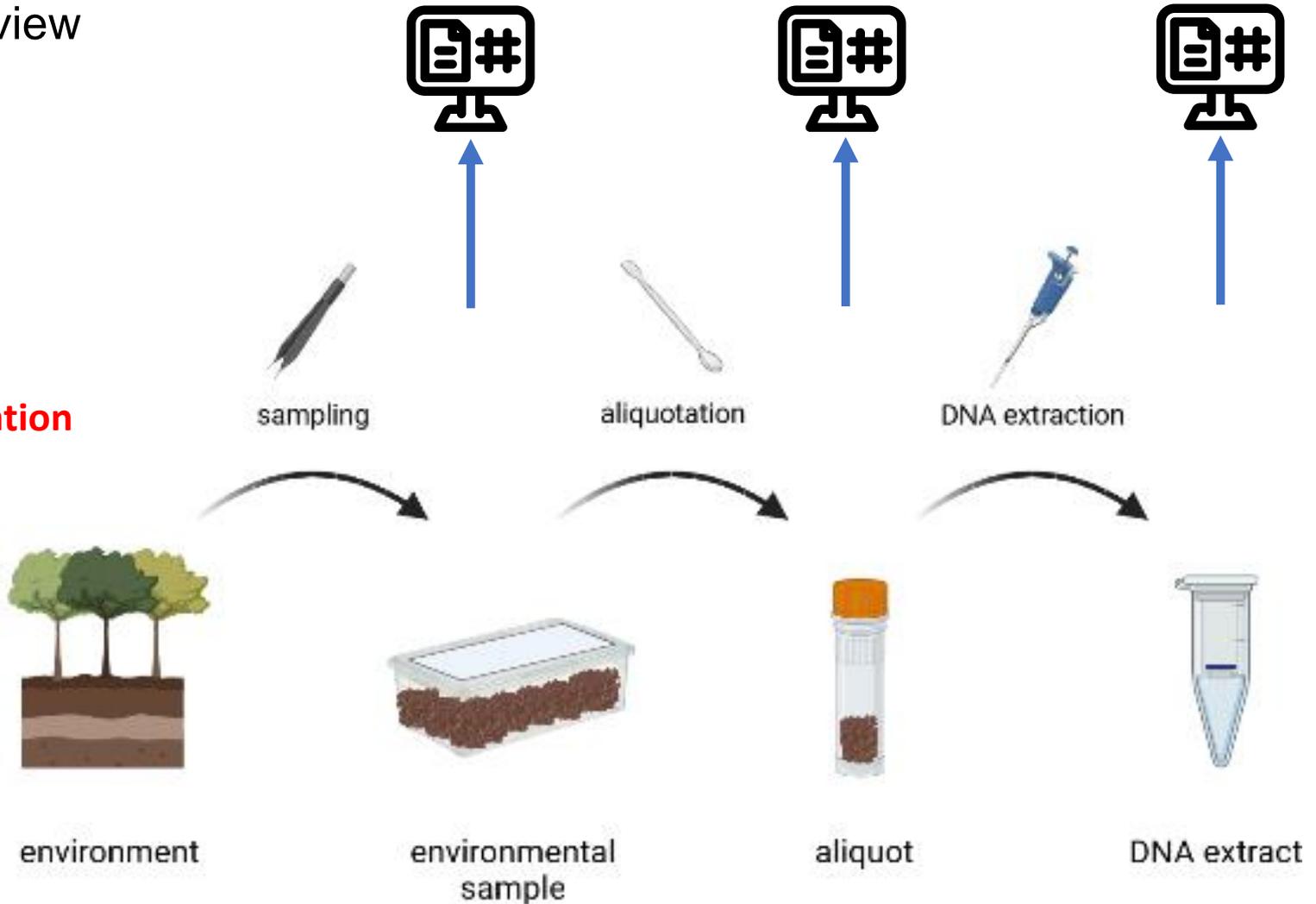
Research and Teaching Activities

- Metabarcoding Workflow Overview

Data

Transformation

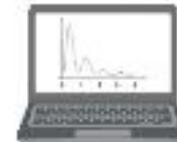
Object



Research and Teaching Activities

- Metabarcoding Workflow Overview

Data



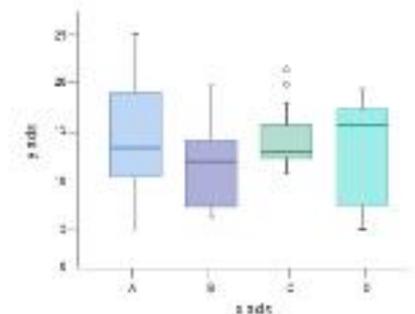
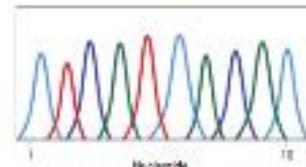
Transformation

PCR amplification

sequencing

data analysis

Object



DNA extract

amplicons

sequences

figures

Research and Teaching Activities

- Research objects
 - Siaya township
 - altitude 1200–1300 m ASL
 - dump sites
 - different types of plastic
 - Mycobiomes on/in
 - plastic particles
 - surrounding soil



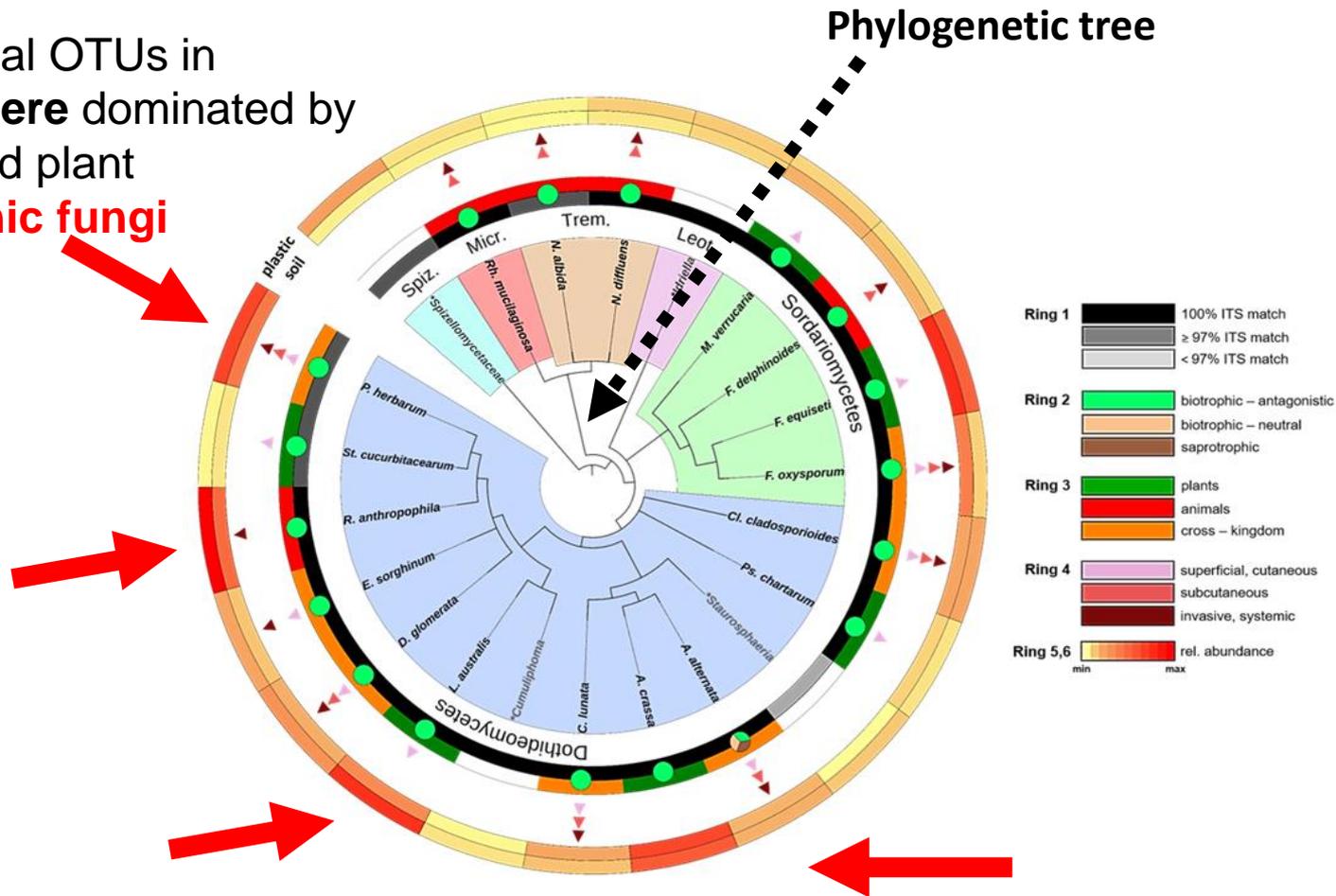
Research and Teaching Activities

- Mycobiomes

- Findings

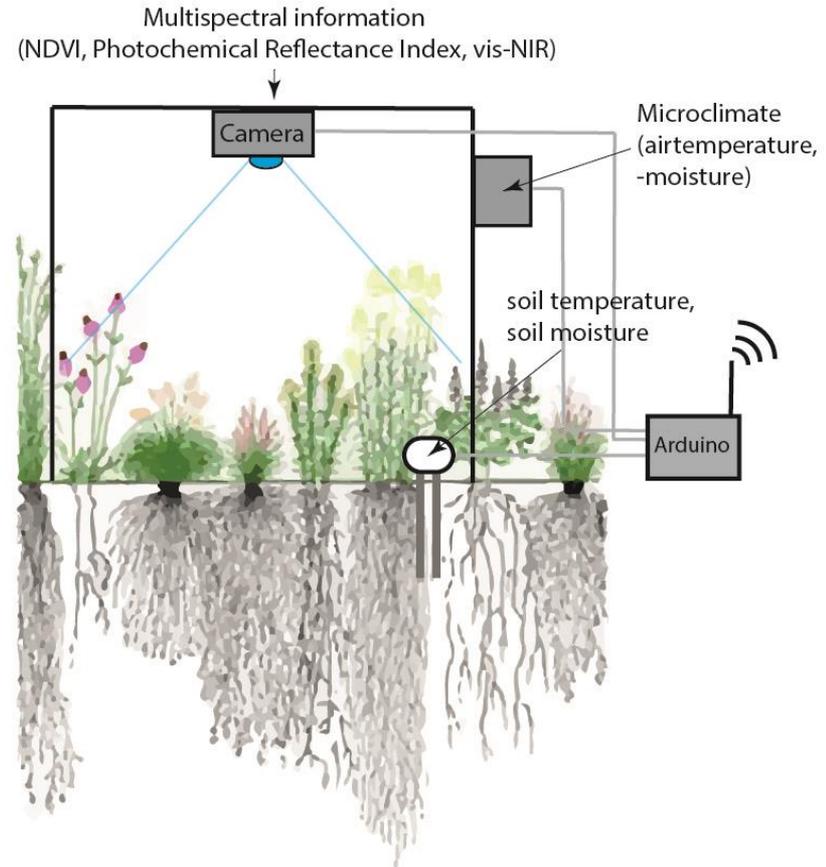
- Most fungal OTUs in **plastisphere** dominated by animal and plant

pathogenic fungi



Research and Teaching Activities

- Grassland Maintenance Experiments
 - Monitoring Station
 - Environmental conditions (soil and atmosphere)
 - Vegetation productivity
 - Alpha diversity of plant and animal species



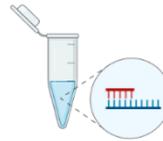
ITCER as **Data** Producer



Monitoring Data Basics

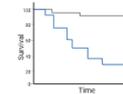
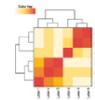
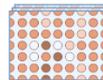
- **Physical objects**

- Environmental samples
- DNA extracts, amplicons
- DNA libraries
- Sec. metabolite extracts



- **Digital objects**

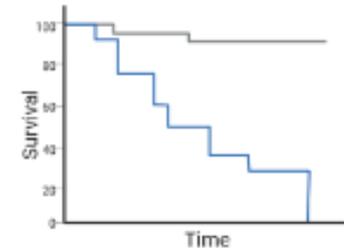
- Streaming data
- Images
- Analysis data (matrices, diagrams)
- DNA libraries (digital)
- Parametrisation data



Monitoring Data Basics

- **Variables**

- Measurement and Contextual data



- **Invariables**

- Design and measurement parametrisation



Monitoring Data Basics

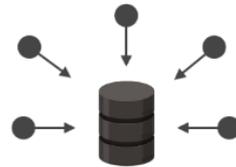
- **Analysis data**

- From measurement data ('variables') and contextual data ('metadata')

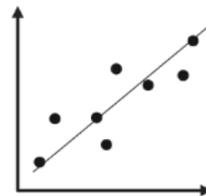


- Transformation

- Aggregation

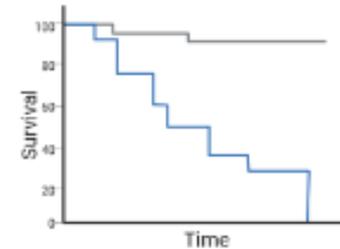


- Correlation



Monitoring Data Basics

- Measurement data (variables)
 - Digital photo of specimen
- Contextual data (variables)
 - Lat / lon coordinates + Date / time
- Analysis data
 - E.g. Measurement data × Contextual data

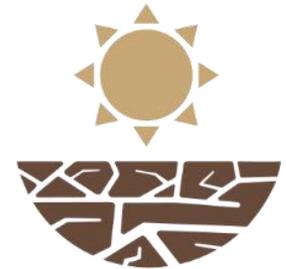
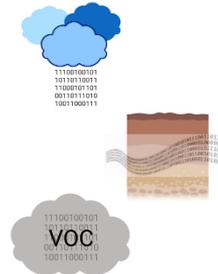


	A	B	C
Cnidaria	200	233	177
Hydra	34	23	14
Cnidophores	7	2	9

Monitoring Data Basics

- **Abiotic factor data**

- Meteorological streaming data
- Soil streaming data
- Volatile compound data



Monitoring Data Basics

- **Biotic factor data**

- Photographical data
- Vegetation analysis (traditional) data
- Community (meta)barcoding data





ITCER:

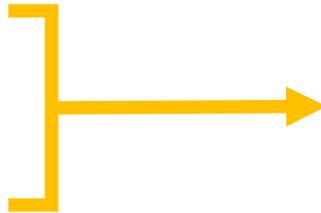
- Biodiversity Site
- Laboratory for Analyses
- Data Node
- Outreach Center and
Communication Platform



→ Research and Teaching

Research and Teaching Activities

- Gathering Contextual Data
- Monitoring
 - Vegetation
 - Insects
 - Mycobiomes
- Grassland Maintenance Experiments



- Monitoring
 - **Vegetation data**
 - Image
 - Logging and streaming
 - Species barcoding
 - Metabarcoding
 - **Insect data**
 - Images
 - Species barcoding
 - Metabarcoding
 - **Microbes and metabolite data**
 - Images
 - Species barcoding
 - Metabarcoding
 - Secondary metabolite screening

Data Products

- Alpha diversity of organisms on standardised plots
- Storage and transfer of image data



DM_DSC09439.JPG_2022-09-24.jp
g



DM_DSC09442.JPG_2022-09-24.jp
g



DM_DSC09465.JPG_2022-09-24.jp
g



DM_DSC09468.JPG_2022-09-24.jp
g



DM_DSC09488.JPG_2022-09-26.jp
g



DM_DSC09489.JPG_2022-09-26.jp
g



DM_DSC09510.JPG_2022-09-26.jp
g



DM_DSC09527.JPG_2022-09-26.jp
g



DM_DSC09537.JPG_2022-09-26.jp
g



DM_DSC09543.JPG_2022-09-26.jp
g



DM_DSC09565.JPG_2022-09-26.jp
g



DM_DSC09567.JPG_2022-09-26.jp
g



DM_DSC09575.JPG_2022-09-26.jp
g



DM_DSC09590.JPG_2022-09-26.jp
g

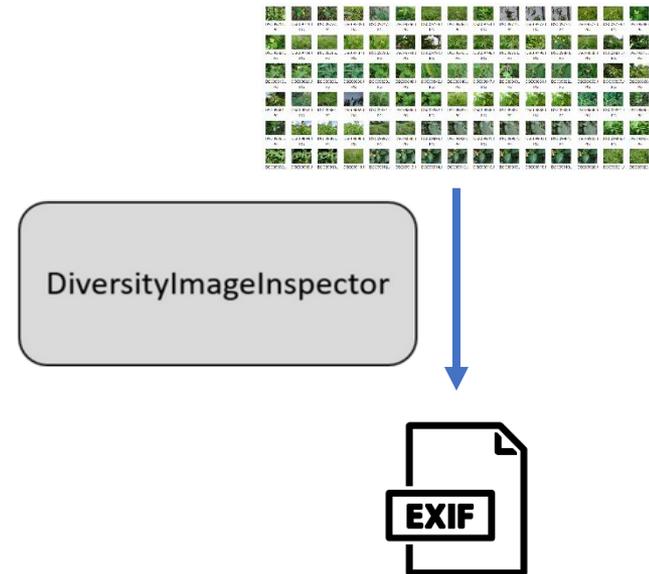


DM_DSC09597.JPG_2022-09-26.jp
g

Data Products

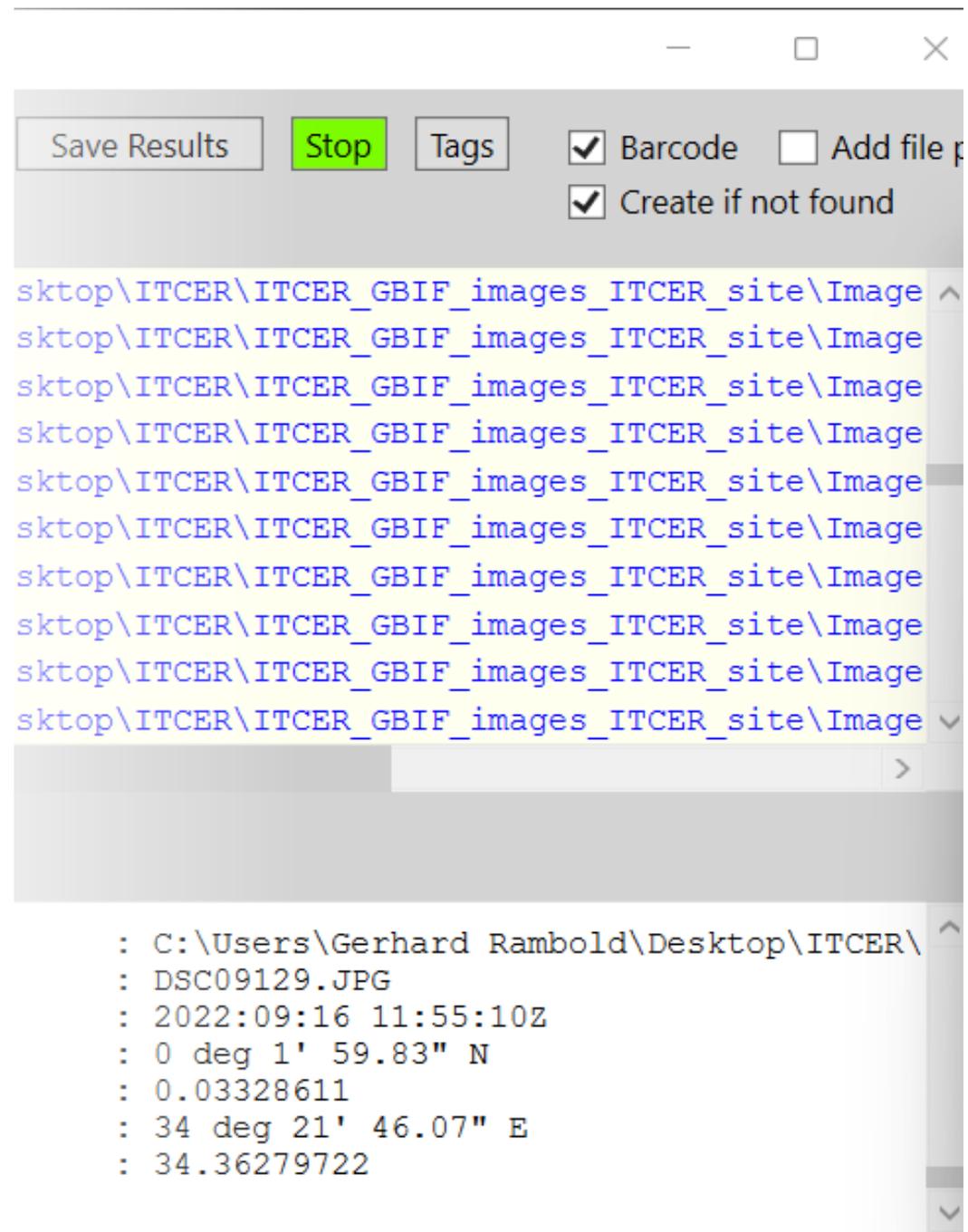


- Organism images
- Image data transfer
 - Diversity Image Inspector
 - EXIF data extraction
 - Diversity Collection
 - EXIF data import
 - Links to images



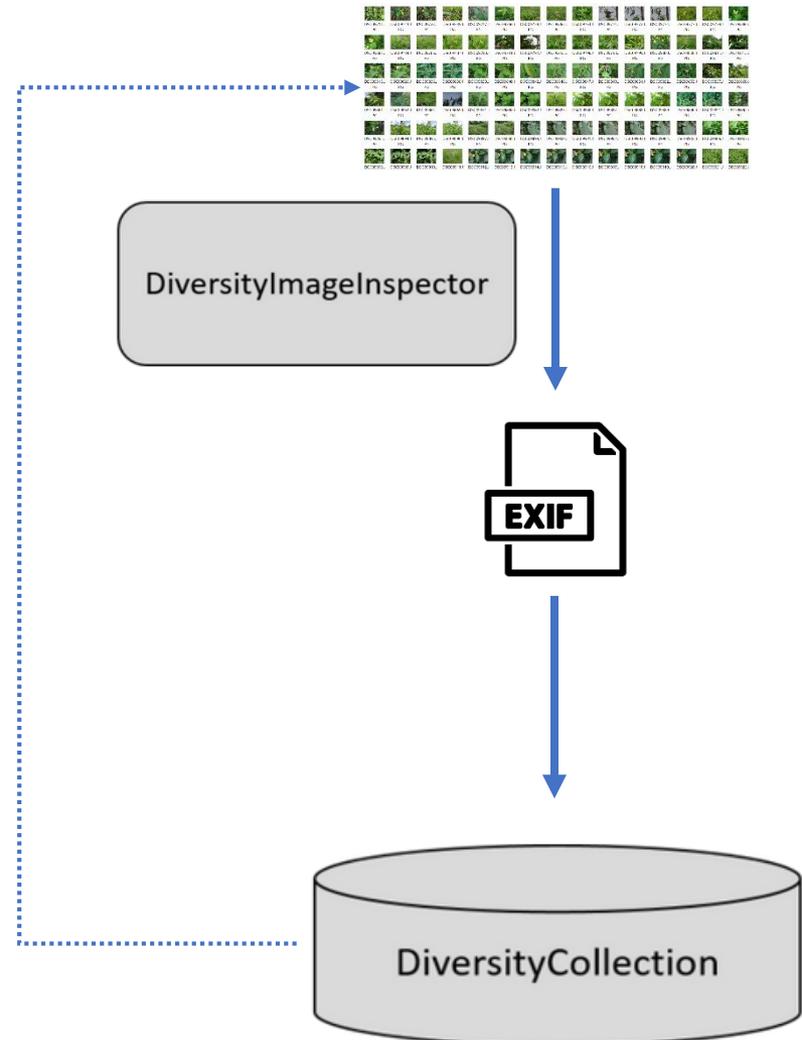
Data Products

- Insect images
 - Diversity Image Inspector
 - EXIF data extraction



Data Products

- Organism images
- Image data transfer
 - Diversity Image Inspector
 - EXIF data extraction
 - Diversity Collection
 - EXIF data import
 - Links to images



Data Products

- Insect images
 - Diversity Image Inspector
 - EXIF data extraction
 - Diversity Collection
 - EXIF data import
 - Links to images

File: C:\Users\Sanz\Documents\Purpurascences\Schmidt-Stohn-Purpurascences Nr. 1-41._150521_fin.txt

Encoding: **ANSI** Start line: 2 End line: 42

First line contains column definition Language / Country: 

Users\Sanz\Documents\Purpurascences\Schmidt-Stohn-Purpurascences Nr. 1-41._150521_fin_20150611_142709.xml

Schedule for import of tab-separated text files into DiversityCollection

Target within DiversityCollection: **Specimen**

Schedule version: 1 Database version: **02.05.58**

Lines: 2 - 42 First line contains column definition:

Encoding: ANSI Language: US

Tables

CollectionEvent (CollectionEvent)

Merge handling: **insert**

Column in table	?	Key	Copy	Pre	Post	File pos.	Transformations	Value	Source	Table
CollectionEventID									Database	
CollectionDay	?					7	Split at Pos. Dir. 1 ⇌		File	
CollectionMonth	?					7	Split at Pos. Dir. 2 ⇌		File	
CollectionYear	?					7	Split at Pos. Dir. 3 ⇌		File	

Specimen	Part 1 StorageLocation Organism 1	Event CountryCache	Event LocalityDescription	Agent 1 CollectorsName Agent 2 CollectorsName	Agent 1 CollectorsNumber	Event CollectionDay Event CollectionMonth Event CollectionYear	Event LocalityDescription
AccessionNumber	LastIdentificationCa Identification 1.1						
M-Nummern	Art	Land	Fundort	Sammler	Koll.-Nr.	Datum	Begleitbäume, O.
M-0275820	Cortinarius purpur...	Ungarn	Szentendre	B. Dima	DB3362	16.10.2008	Quercus petraea
M-0275821	Cortinarius purpur...	Ungarn	SzSzB, Bátorliget	L. Albert & B. Dima	DB1339	02.10.2004	Quercus robur
M-0275822	Cortinarius purpur...	Deutschland	Kyffhäuser/Leop...	G. Schmidt-Stohn	SSt07-211	09.10.2007	Quercus
M-0275823	Cortinarius subpu...	Italien	Matschatsch	C. Rossi	Rossi4034	03.10.1998	Fagus
M-0275824	Cortinarius purpur...	Italien	S. Vito Comedo	C. Rossi	Rossi4906	22.10.2000	Picea, Castanea
M-0275825	Cortinarius purpur...	Italien	Calangianus	C. Rossi	Rossi5053	02.11.2000	Quercus suber, G

Monitoring Data Basics

- Legal issues

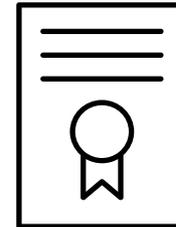
- Copyright



- Author's rights or intellectual property

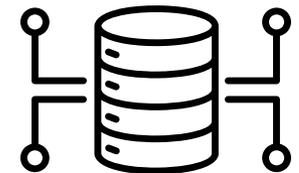
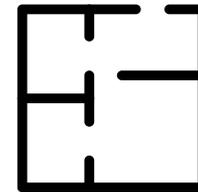
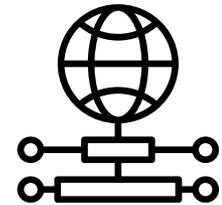
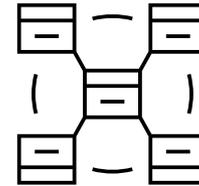


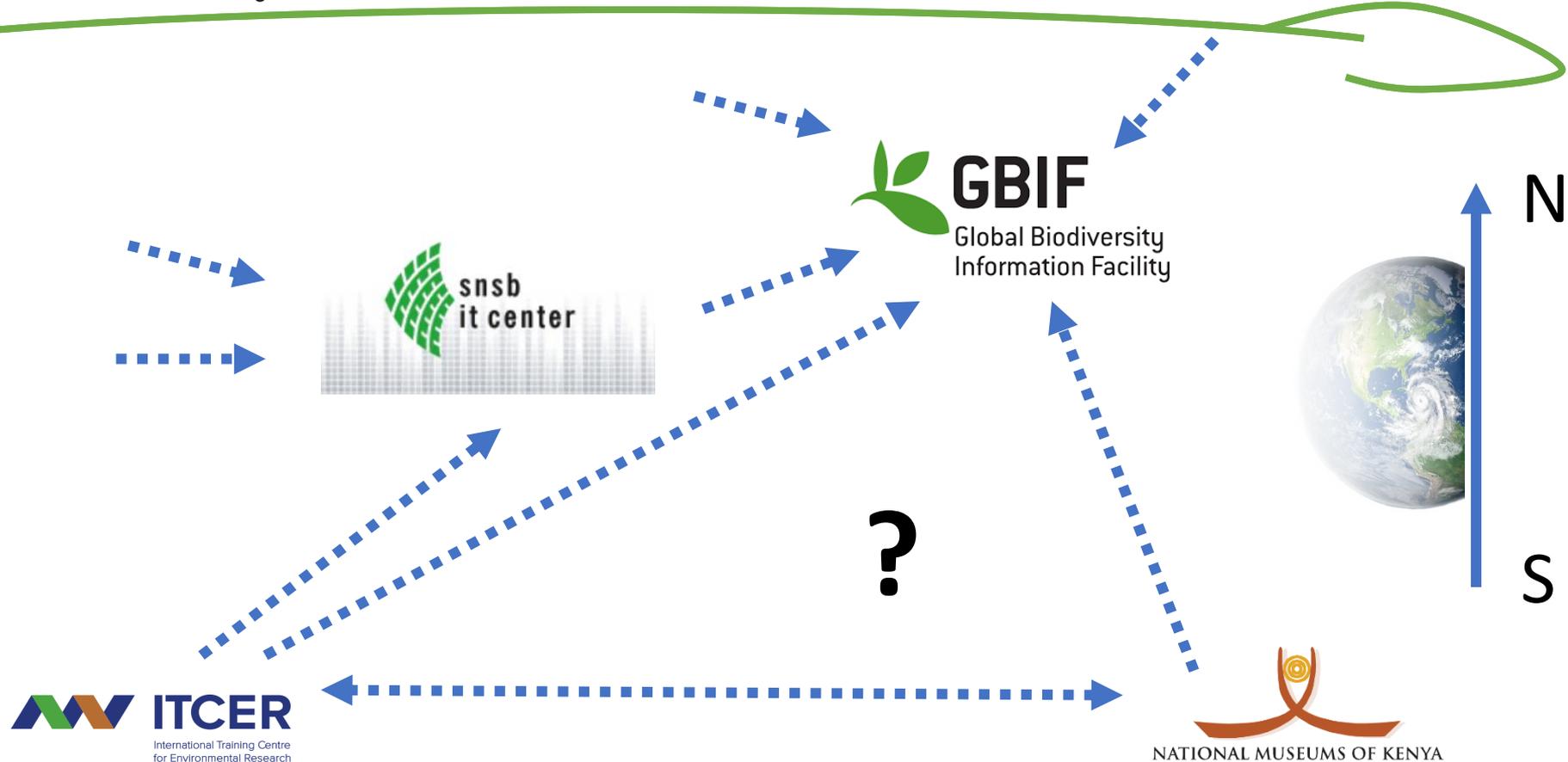
- Permits



Monitoring Data Basics

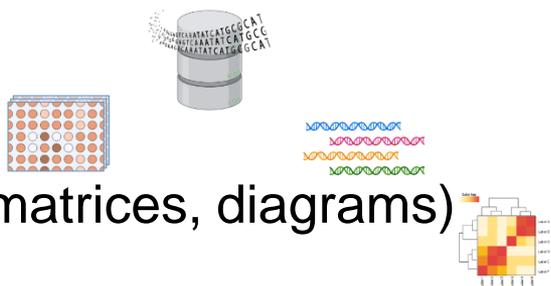
- Data flow
 - Schema
 - Protocol
 - Physical and digital objects location
 - Plots
 - Databases
 - Image repositories





• **Digital objects**

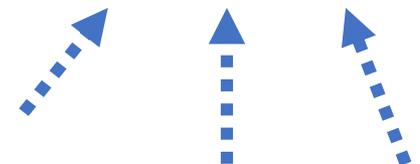
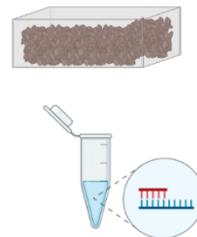
- Streaming data
- Images
- Analysis data (matrices, diagrams)





- **Physical objects**

- Environmental samples
- DNA extracts, amplicons
- DNA libraries
- Sec. metabolite extracts



ITCER Workshop, Oct 27th to 28th 2022

Expert Workshop on GBIF Data

ITCER Research, Training and Data

