ITCER in GBIF: To become a publisher of occurrence data

Dagmar Triebel

Expert Workshop on GBIF Data
Publication in Africa, 17.10.2024

GBIF network

Datasets
106,289

Hosted portals
21

Country Participants
64

Peer-review papers using data
10,715

Organizational Participants
43

Average records downloaded per month (2024)
201.5 billion

Publishers
2,249

Species occurrence records
2,973,669,729

GBIF Participant countries: 142

1 July 2024

142 countries/areas with institutions
sharing data through GBIF

Voting Participants

Associate Participants

https://www.gbif.org/the-gbif-network
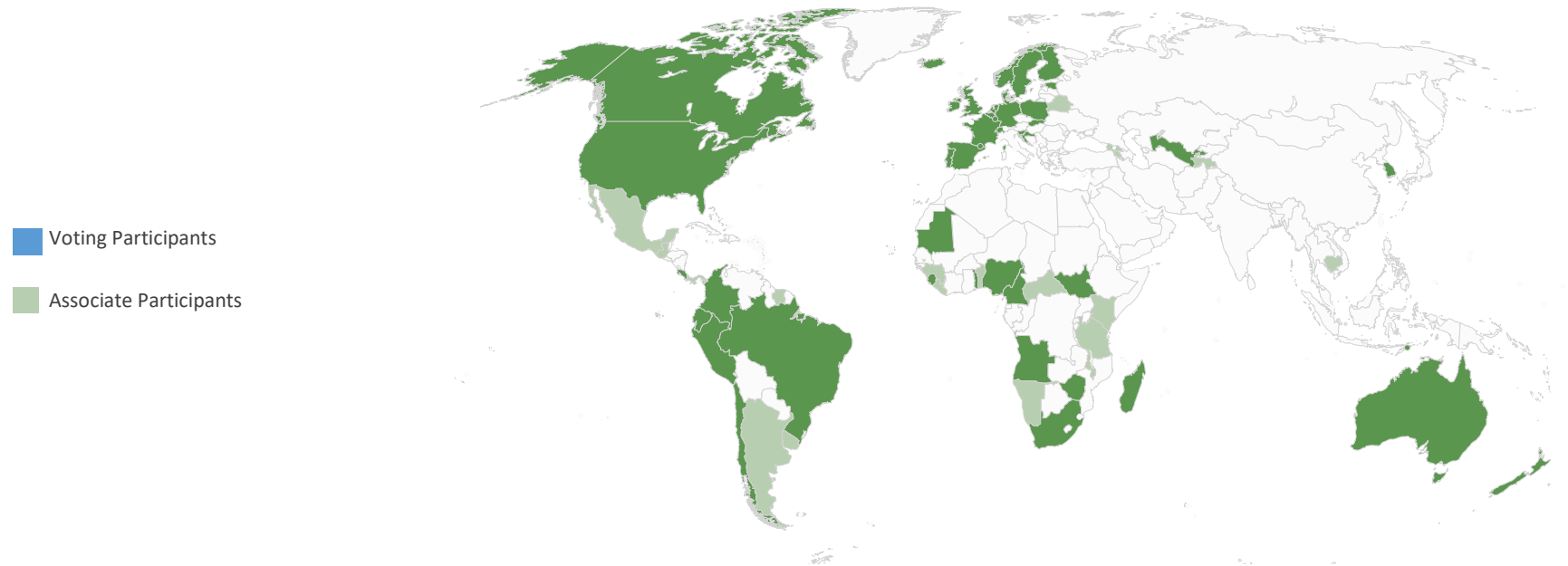
GBIF data publishing institutions: 2,249

First 11 of 142 countries/areas with 2,249 institutions sharing data through GBIF

Country	Institutions
United States	342
Colombia	227
United Kingdom	170
Australia	120
Spain	117
Brazil	114
Russian Federation	98
France	61
Ecuador	54
Canada	48
Netherlands	48

Number of active publishers by country or area

0   1   2   4   10   40   100+
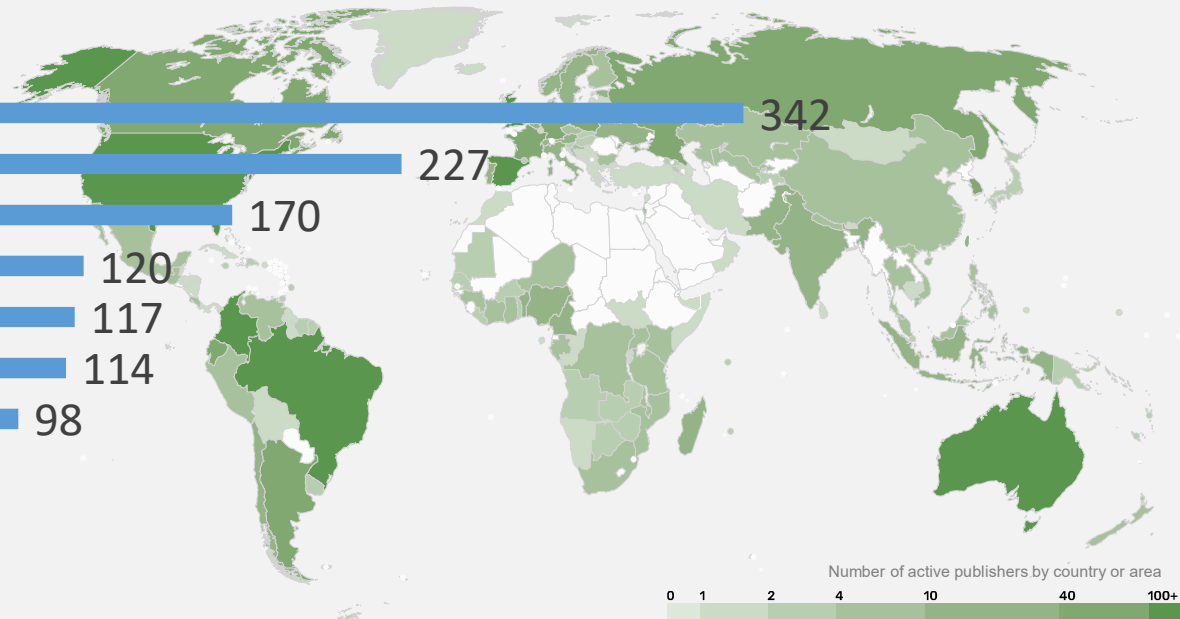
GBIF Network
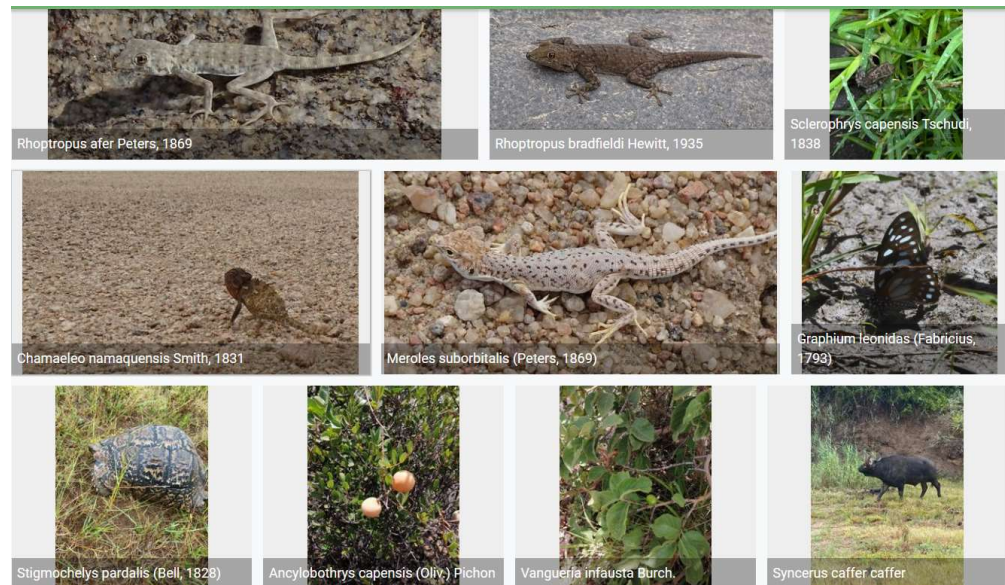## Species occurrence records with multimedia evidence

180.7 million records with
taxonomically identified images
- 120.6 million human observations
- 55.0 million specimens
- 2.8 million material samples
- 1.4 million fossil specimens

1,420,991 audio files
10,418 videos

https://www.gbif.org/occurrence/gallery

GBIF Network: www.gbif.org

- GBIF Secretariat, IT Network

- GBIF Participants, e.g. nodes of countries, e.g. of Kenya

- GBIF Data Publishers (organisations), e.g. ITCER

- GBIF Datasets with occurrence data records, e.g. ITCER Macrobiome dataset on site

ITCER in GBIF – a guideline

- To become a data publishing institution

    Fulfill administrative endorsement criteria

    Fulfill technical endorsement criteria (agreement with hosting centres or
    participant node organisation)

- After that: ready to publish one or more datasets

- Attention! No individuals can provide data for GBIF

- GBIF Data Publishers are organisations.

Expert Workshop on GBIF Data Publication in Africa, 17.10.2024

ITCER in GBIF – a guideline

5. **Data Publishing**: The process of and agreements for making data freely and universally available on the Internet.

6. **Data Publisher**: A custodian that makes data publicly available, whether or not they are the data owner. In cases where data publishers are not the data owners, they will have declared that they have permission to make the data available under terms of the GBIF Data Publisher Agreement.

ITCER in GBIF –To become a data publishing institution

- Data management and publication infrastructure might be hosted by SNSB like for Komorov Bot. Inst. (ABCD/ BioCASe option with long-term preservation via DWB services)

ITCER in GBIF –To become a data publishing institution

- Data publication infrastructure might be hosted by National Museums of Kenya like for Strathmore University (IPT option)

ITCER in GBIF –To become a data publishing institution

- endorsed by GBIF Kenya (for both options: ABCD/ BioCAse and IPT)

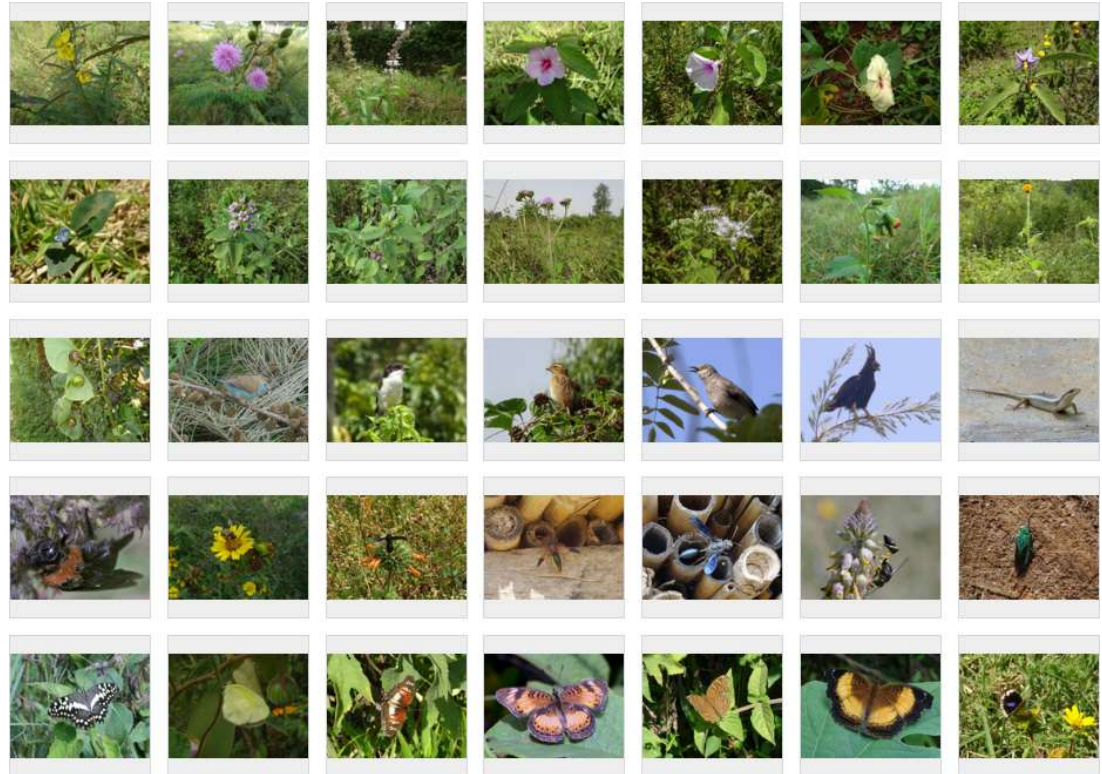# ITCER – GBIF occurrence dataset as a pilot

- Title, like „Photomonitoring of Macroorganisms occurring on ITCER site and their interactions"

- Scope, agents, licences

- Data pipelines

- Relation to GBF data paper

- *…. to be discussed in the afternoon*

snsb it center

SNSB Staatliche Naturwissenschaftliche Sammlungen Bayerns

ITCER Tools for GBIF dataflow?

Rukaya Johaadien, head engineer from GBIF Norway and winner of the first prize entry in the 2024 Ebbe Nielsen Challenge. Photo courtesy of Ms Johaadien.

ChatIPT, a chatbot that cleans and standardizes spreadsheets, creates basic metadata and guides students and researchers through the process of publishing data into the GBIF network, has won GBIF's 2024 Ebbe Nielsen Challenge.
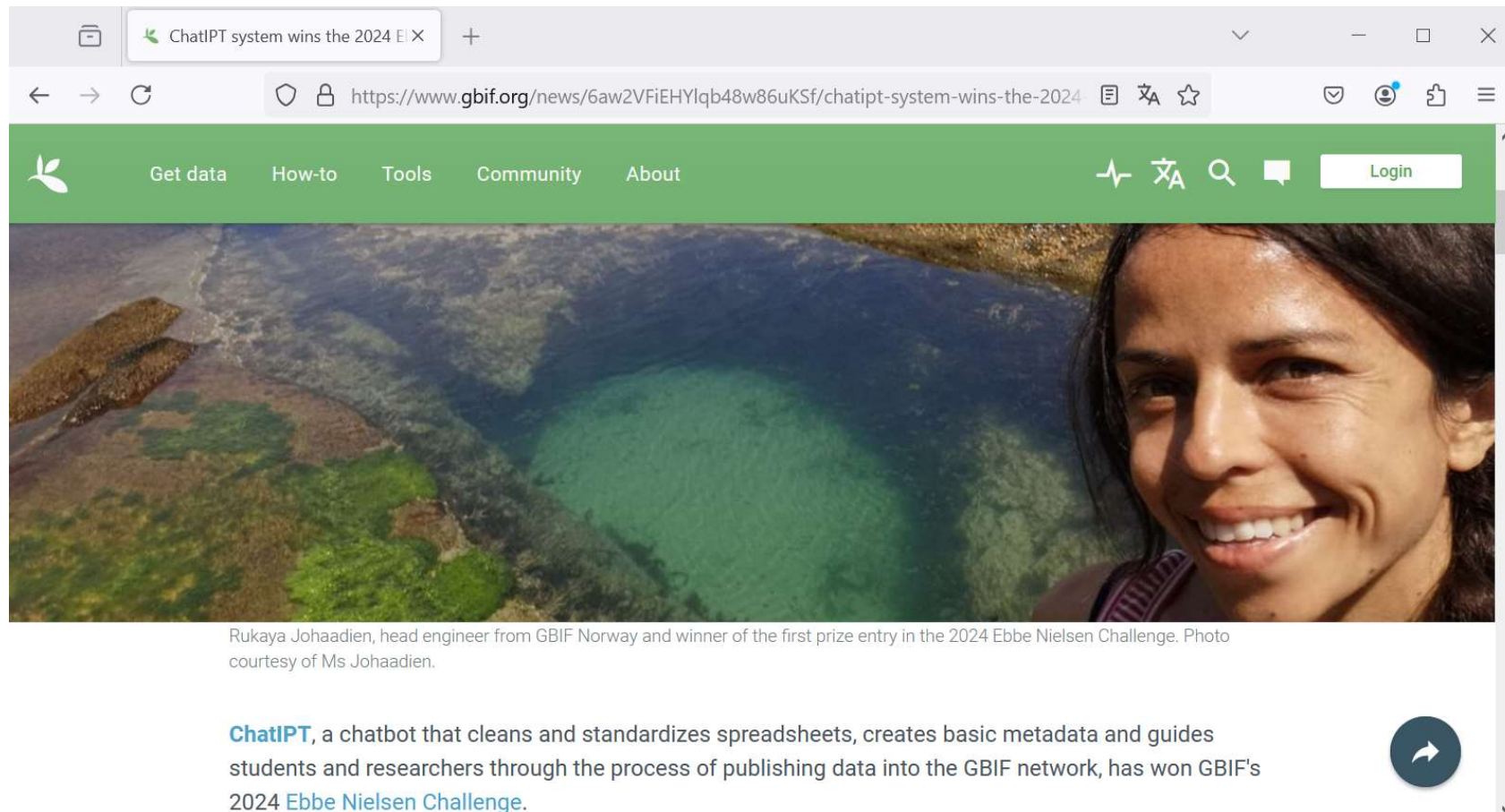
Expert Workshop on GBIF Data Publication in Africa, 17.10.2024

GBIF material linked under:

http://www.snsb.info/SNSBwiki/Wiki.jsp?page=Expert-Workshop-on-GBIF-Data-Publication-in-Africa-2024

Relevant GBIF Resources

GBIF overview: PowerPoint slides↗

GBIF Africa Network↗

GBIF Participant Node Kenya↗

GBIF Kenya Activity Report 2023↗

GBIF Data Publisher Endorsement Guidelines↗ (with involvement of the GBIF Node Kenya)

Become a GBIF Data Publisher↗ (to be accepted by ITCER, endorsed by the GBIF Node Kenya)

GBIF data publisher agreement↗

Data Hosting Centres↗ (to be accepted by ITCER and SNSB - BioC/

Agreement on GBIF Hosted Repository Service↗ (Repository Owne

Blog on Choose dataset class↗:

- ○ example for class occurrence data↗
  - ○ example for class checklist data↗, see data quality requirements↗
  - ○ example for class sampling event data↗

Blog on GBIF Occurrence License Processing↗

Blog on Sharing images, sounds and videos on GBIF↗

GBIF occurrence dataset from Russia hosted by SNSB data publisher↗ as example

snsb it center

SNSB Staatliche Naturwissenschaftliche Sammlungen Bayerns

Thank you!

Screenshots and resources mentioned in the slides are cited under

http://www.snsb.info/SNSBwiki/Wiki.jsp?page=Expert-Workshop-on-GBIF-Data-Publication-in-Africa-2024