

Concepts for service integration and data publication via NFDI4Biodiversity and GBIF

Tanja Weibulat, Dagmar Triebel

45. Diversity Workbench Workshop
ITCER, Kenya, 09.10.2023

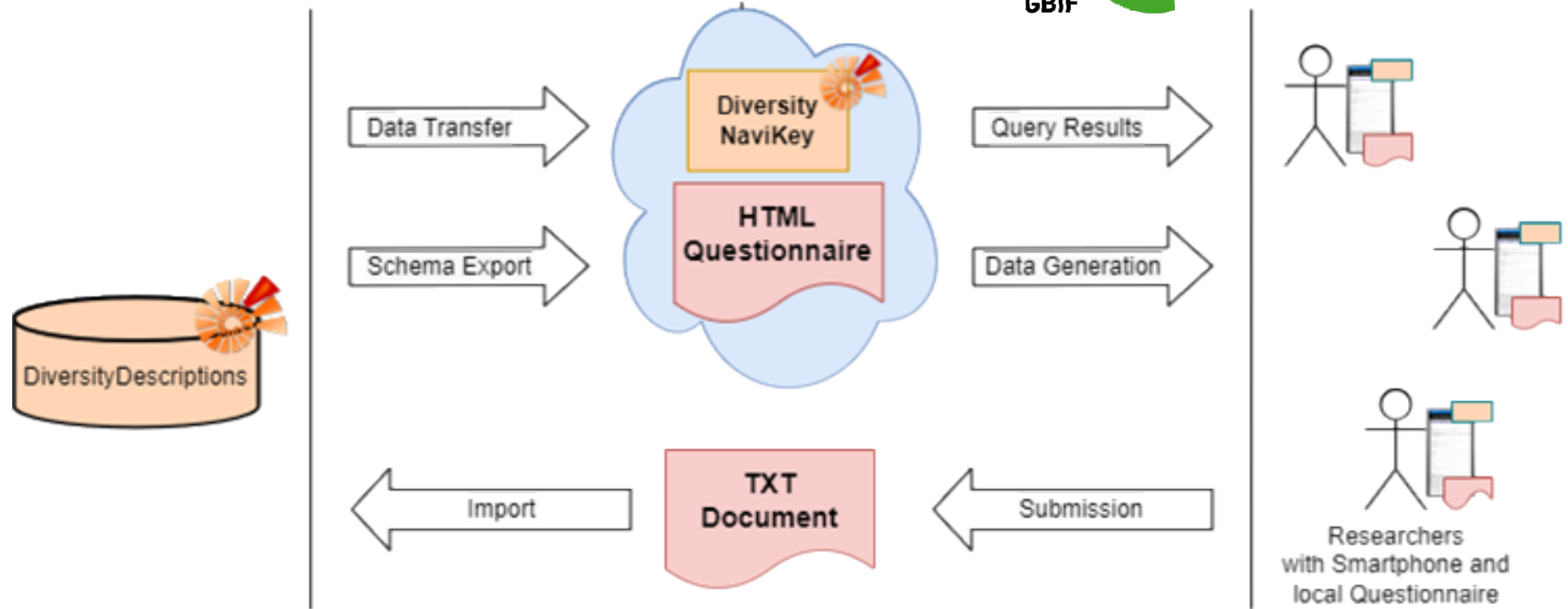


EthnoMycAfrica data pipeline

NFDI 4
BIODIVERSITY
BIODIVERSITY, ECOLOGY & ENVIRONMENTAL DATA



NFDI-RDC
Digital Objects

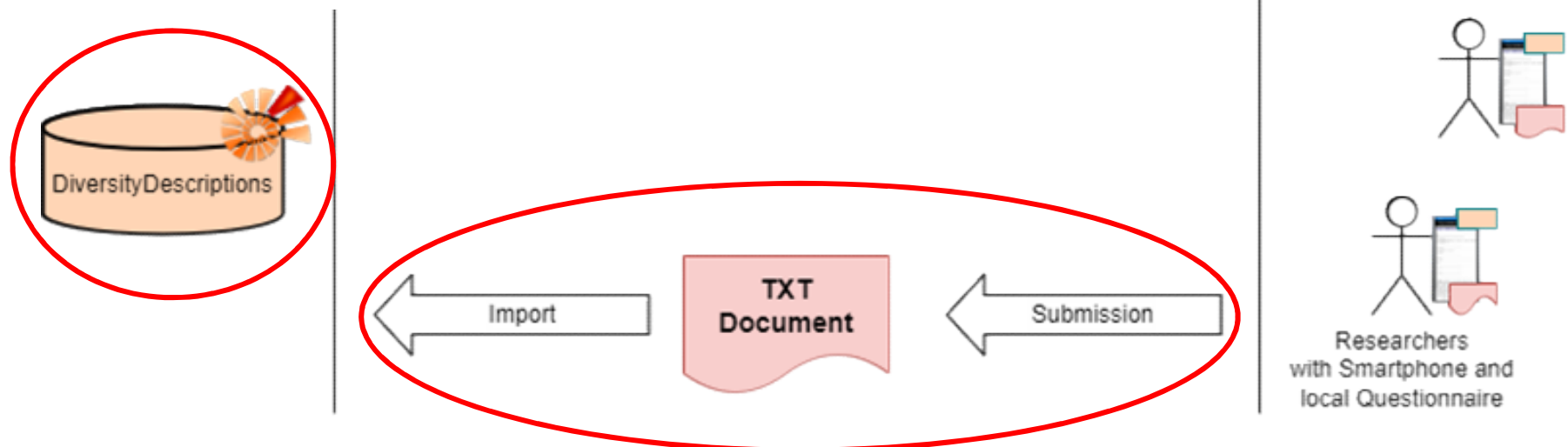


Kinge et al. 2023, Fig. 6 (modified) <https://doi.org/10.1017/S0024282923000294>




Managing and storing services realised

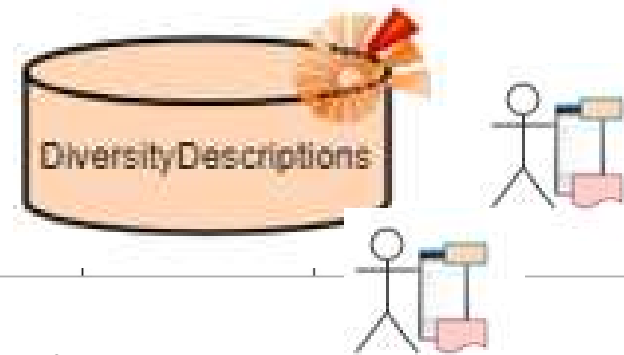
- Data of the EthnoMycAfrica project are managed in a DiversityDescriptions installation.
- The installation is administrated at SNSB.
- The EthnoMycAfrica project team has access to the data for management and curation purposes.



Kinge et al. 2023, Fig. 6 (modified) <https://doi.org/10.1017/S0024282923000294>

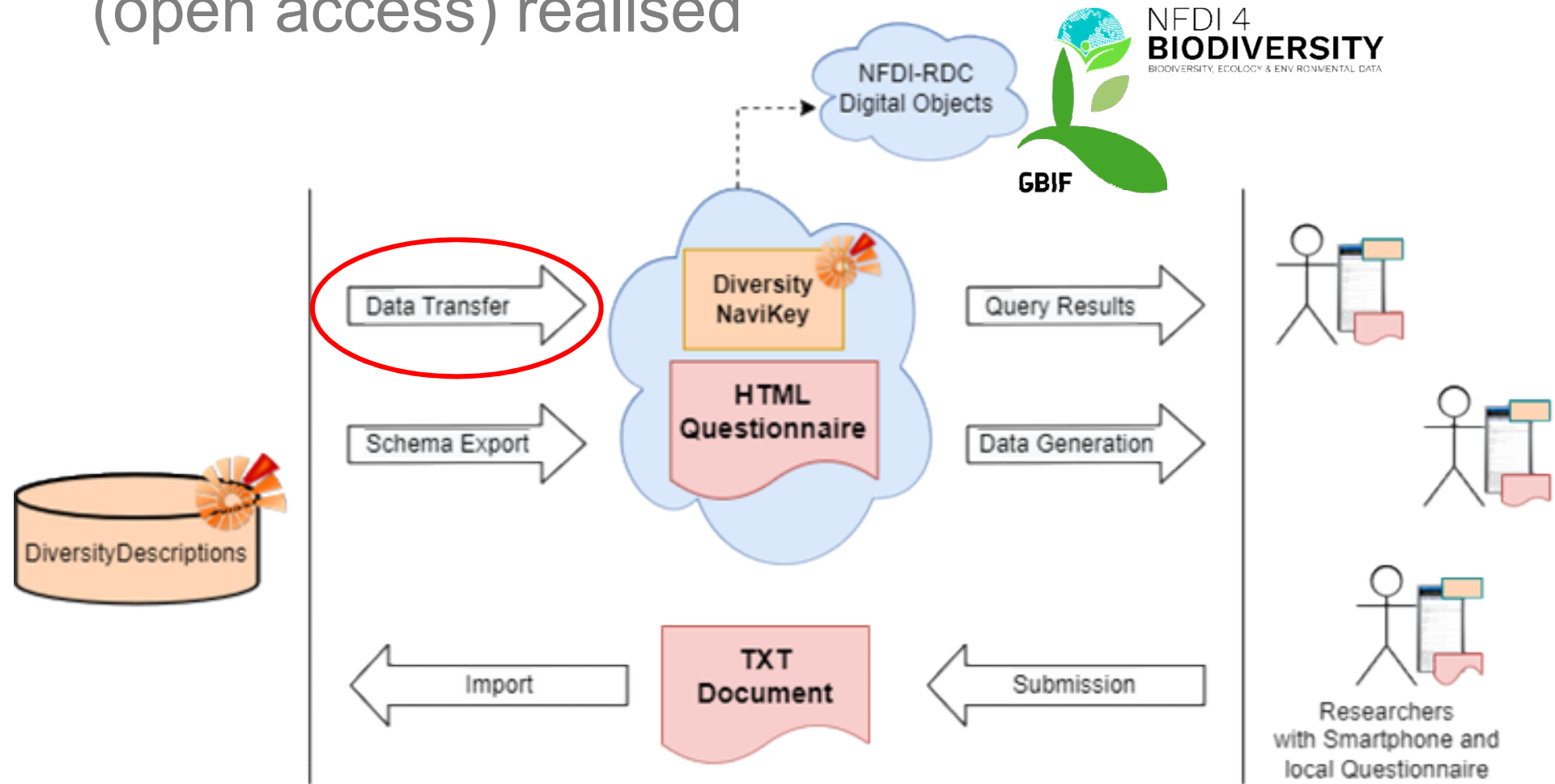
Managing and storing services realised

DiversityNaviKey code and project logo	Title and description of the data source (dataset)	Responsible/ author(s)	Creative Commons Licenses (CC)  (textual data and hosted images)	Number of items, number of item values (*)	Number of descriptors, number of categorical descriptor states (-)	Number of linked images and documents hosted by SNSB	Linked internet resources	Access point with citation (e.g. re3data, DOI)	References and notes
UBTMYCethnomycdesc / EthnoMycAfrica (....)	EthnoMycAfrica – Ethnomycological knowledge in Africa	Kinge, Rosemary Tonjock; Jefwa, Joyce Mnyazi; Houdanon, Roel D.; Kamalebo, Héritier Milenge; Abduzalem, Ahmed; Gryzenhout, Marieka; Rambold, Gerhard	CC BY	1,366 items with 13,490 categorical state values and 6,438 text values	34 descriptors with 21 categorical descriptors (486 categorical states), 13 free text descriptors	-	-	-	item status from Mar. 2023 (DD); HTML Questionnaire for EthnoMycAfrica Database , to run the HTML questionnaire locally please download the zip-archive



https://diversityworkbench.net/Portal/DWB_DiversityNaviKey_data_sources

Processing service to feed the DNK interface (open access) realised



Kinge et al. 2023, Fig. 6 (modified) <https://doi.org/10.1017/S0024282923000294>

Processing service to publish SDD-EML structured data products with DOI as NFDI4Biodiversity service planned



Staatliche Naturwissenschaftliche Sammlungen Bayerns

Info Settings Prefixes **DOIs**

Create DOI

Export DOI Metadata

State

- Findable 6
- Draft 1

Resource Type

- Dataset 7

Year created

Type to search, For example 10.4121/17185607.v1

Search

Reset All

7 DOIs

Sort by Date Updated

Mobilisation of data for GFBio services - Belmont Forum Open Data Survey 2014 Use Case Dataset

Birgit Schmidt, Birgit Gemeinholzer, Andrew Treloar, Jonathan Hodge, André Santanchè, Kim Oakley, Veronica Sanz, Dagmar Triebel, EML published 2020 via Staatliche Naturwissenschaftliche Sammlungen Bayerns – SNSB IT Center, München

The data collection is that from the original data publication, i.e. Schmidt, B., Gemeinholzer, B., Treloar, A., Hodge, J., Santanchè, A., & Oakley, K. (2015). Belmont Forum Open Data Survey 2014 [Data set]. Zenodo. <http://doi.org/10.5281/zenodo.16384>. The Belmont Forum Open Data Survey was from September -- November 2014, organised by the Open Data working group of the E-Infrastructure and Data Management Collaborative Research Action (<https://www.belmontforum.org/>). The findings of the Belmont Forum's survey on Open Data were discussed by Schmidt, B., Gemeinholzer, B., Treloar, A. (2016): Open Data in Global Environmental Research: The

<https://doi.datacite.org/repositories/zbmed.snsb>

Tanja Weibulat, 45. Diversity Workbench Workshop, ITCER, Kenya, 09.10.2023



Processing service to publish SDD-EML structured data products with DOI as NFDI4Biodiversity service planned

The screenshot shows the GfBio search interface. At the top left is the GfBio logo. The top right contains navigation links: GfBio Services, GfBio e.V., and NFDI 4 Bio. A search bar contains the text 'SDD' with 'SEARCH' and 'SEMANTIC SEARCH' buttons. Below the search bar is a 'Dataset Basket' showing '0' items. The results section indicates '4 entries found' with 'Previous', '1', and 'Next' navigation buttons. Two search results are visible, each with a '2020' tag, 'Open Access' status, and 'SNSB' affiliation. The first result is by Schmidt, Birgit; Gemeinholzer, Birgit; Treloar, Andrew; Hodge, Jonathan; Santanchè, André; Oakley, Kim; Sanz, Veronica; Triebel, Dagmar (2020): Mobilisation of data for GfBio services - Belmont Forum Open Data Survey 2014 Use Case. The second result is by Söhngen, Carola; Bunk, Boyke; Podstawka, Adam; Gleim, Dorothea; Overmann, Jörg (2020): Mobilisation of data for GfBio services - Bacterial Descriptions Use Case. A filters sidebar on the left shows 'Data Center SNSB' and 'Data Center' options. The bottom right of the screenshot features the NFDI 4 BIODIVERSITY logo and the SNSB IT Center logo.

<https://search.gfbio.org/>

Tanja Weibulat, 45. Diversity Workbench Workshop, ITCER, Kenya, 09.10.2023



Processing service to publish SDD-EML structured data products with DOI as NFDI4Biodiversity service planned



GFBio Data Center & Archive

IT Center of the Staatliche Naturwissenschaftliche Sammlungen Bayerns



Title:

Mobilisation of data for GFBio services - Belmont Forum Open Data Survey 2014 Use Case

Description:

The data collection is that from the original data publication, i.e. Schmidt, B., Gemeinholzer, B., Treloar, A., Hodge, J., Santanchè, A., & Oakley, K. (2015). Belmont Forum Open Data Survey 2014 [Data set]. Zenodo. <http://doi.org/10.5281/zenodo.16384> . The Belmont Forum Open Data Survey was from September – November 2014, organised by the Open Data working group of the E-Infrastructure and Data Management Collaborative Research Action (<https://www.belmontforum.org/>). The findings of the Belmont Forum's survey on Open Data were discussed by Schmidt, B., Gemeinholzer, B., Treloar, A. (2016): Open Data in Global Environmental Research: The Belmont Forum's Open Data Survey. PLOS One, DOI: [10.1371/journal.pone.0146695](https://doi.org/10.1371/journal.pone.0146695) . In 2015, the data were restructured following the EML 2.1.1 standard and the SDD 1.1 standard and used as good practice example to optimize internal pipelines for data processing within the project GFBio (<https://www.gfbio.org/>).

Citation:

Schmidt, Birgit; Gemeinholzer, Birgit; Treloar, Andrew; Hodge, Jonathan; Santanchè, André; Oakley, Kim; Sanz, Veronica; Triebel, Dagmar (2020). Mobilisation of data for GFBio services - Belmont Forum Open Data Survey 2014 Use Case. [Dataset]. Version: 20201125. Staatliche Naturwissenschaftliche Sammlungen Bayerns – SNSB IT Center, München. <https://doi.org/10.25897/5/nhc7-0d72> .

Data:

Project	Archive	hash	Size	Transfer Date	Publication Date	File	Data format	Identifier
BELMONTsurveydesc	21	b3104c702fdffb9064...	115.9 kB	2020-11-25	2020-11-25	Download	EML	DOI: 10.25897/5/j2cs-q186
BELMONTsurveydesc	20	6d0a8c623a0def22dc...	264.4 kB	2020-11-25	2020-11-25	Download	EML-SDD	DOI: 10.25897/5/nhc7-0d72

Keywords:

E-Infrastructures, FAIR, Global Environmental Change, Open Data, Sharing, Structured Survey Data, World

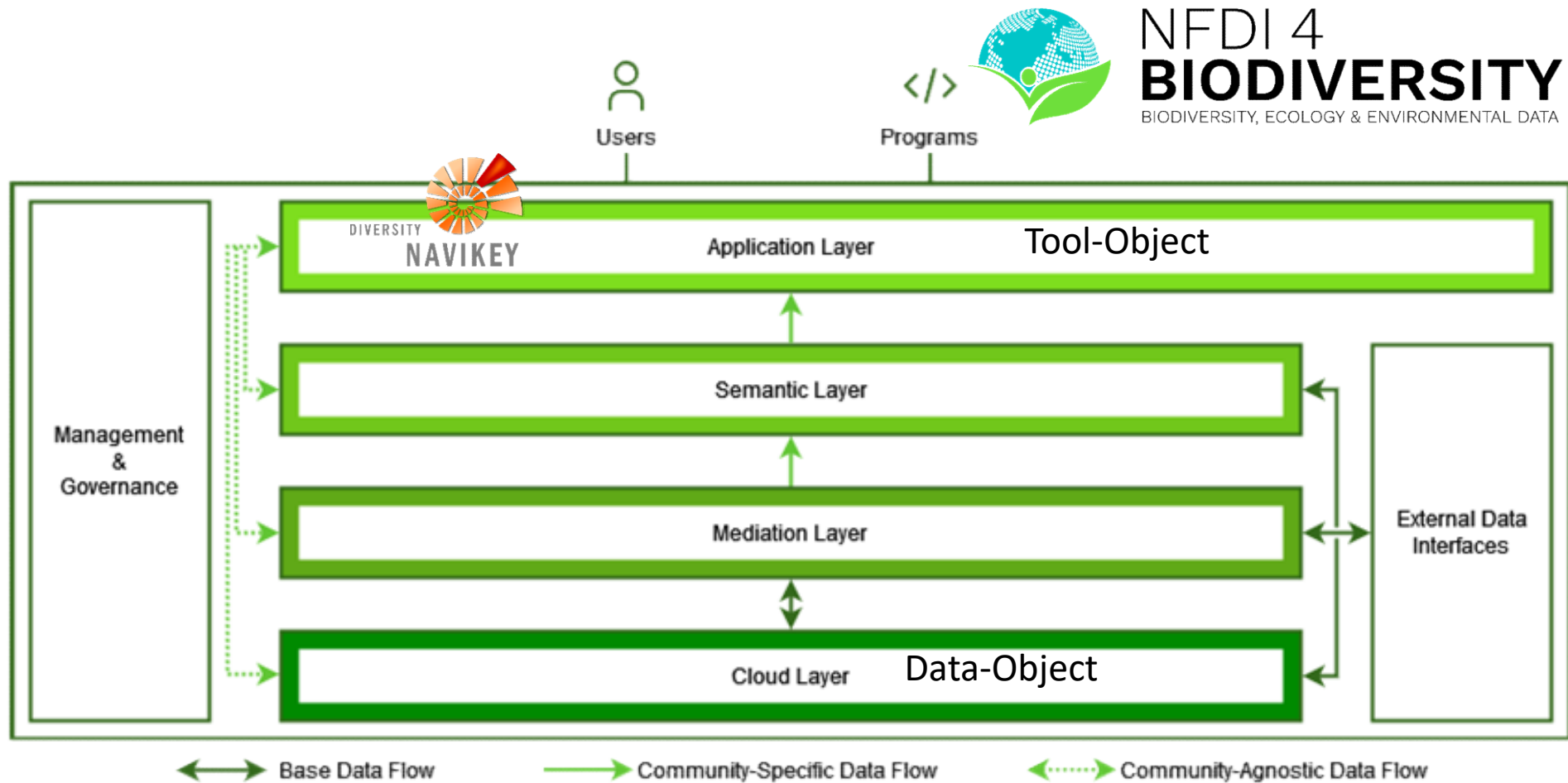
License:

CC BY 4.0 International

https://data.snsb.info/dataset/gfbio/projects_gfbio/968/?archiveid=20



Processing service to integrate data and service in NFDI4Biodiversity RDC cloud planned

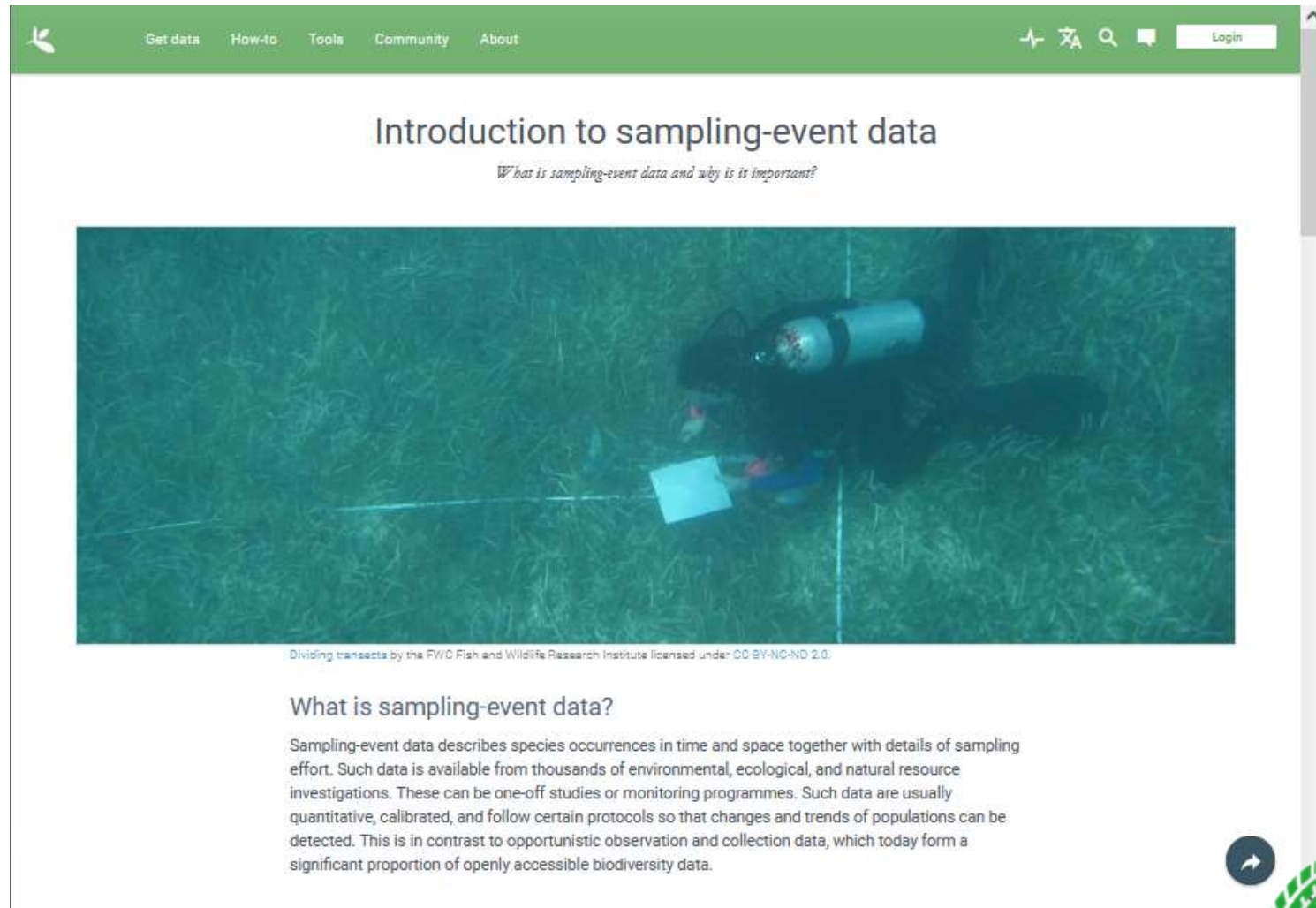


NFDI 4
BIODIVERSITY
BIODIVERSITY, ECOLOGY & ENVIRONMENTAL DATA

<https://www.nfdi4biodiversity.org/de/>



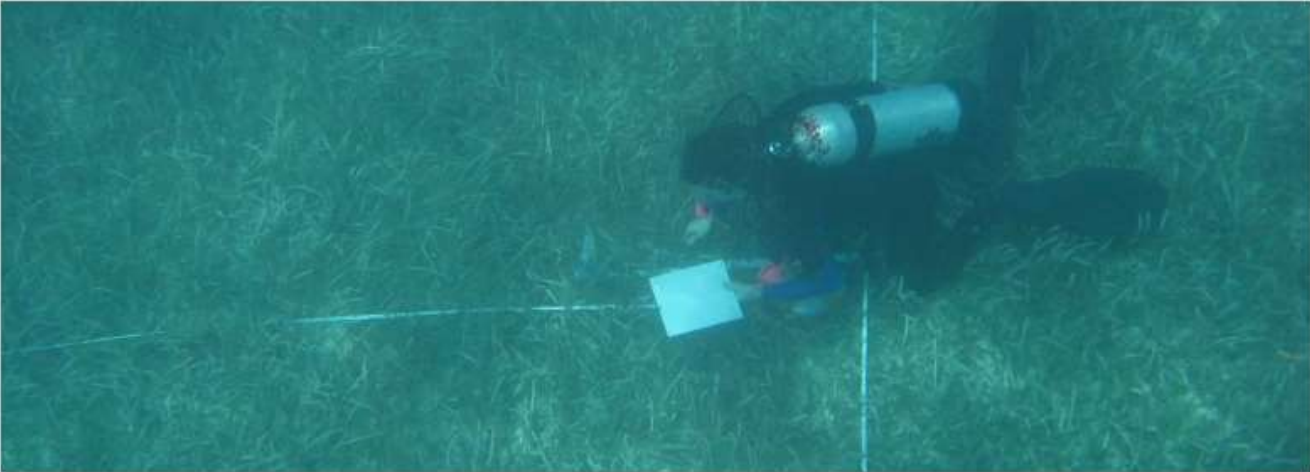
Processing service to publish data products (digital objects) via GBIF Sampling-Event planned



Get data How-to Tools Community About

Introduction to sampling-event data

What is sampling-event data and why is it important?



Dividing transects by the FWC Fish and Wildlife Research Institute licensed under CC BY-NC-ND 2.0.

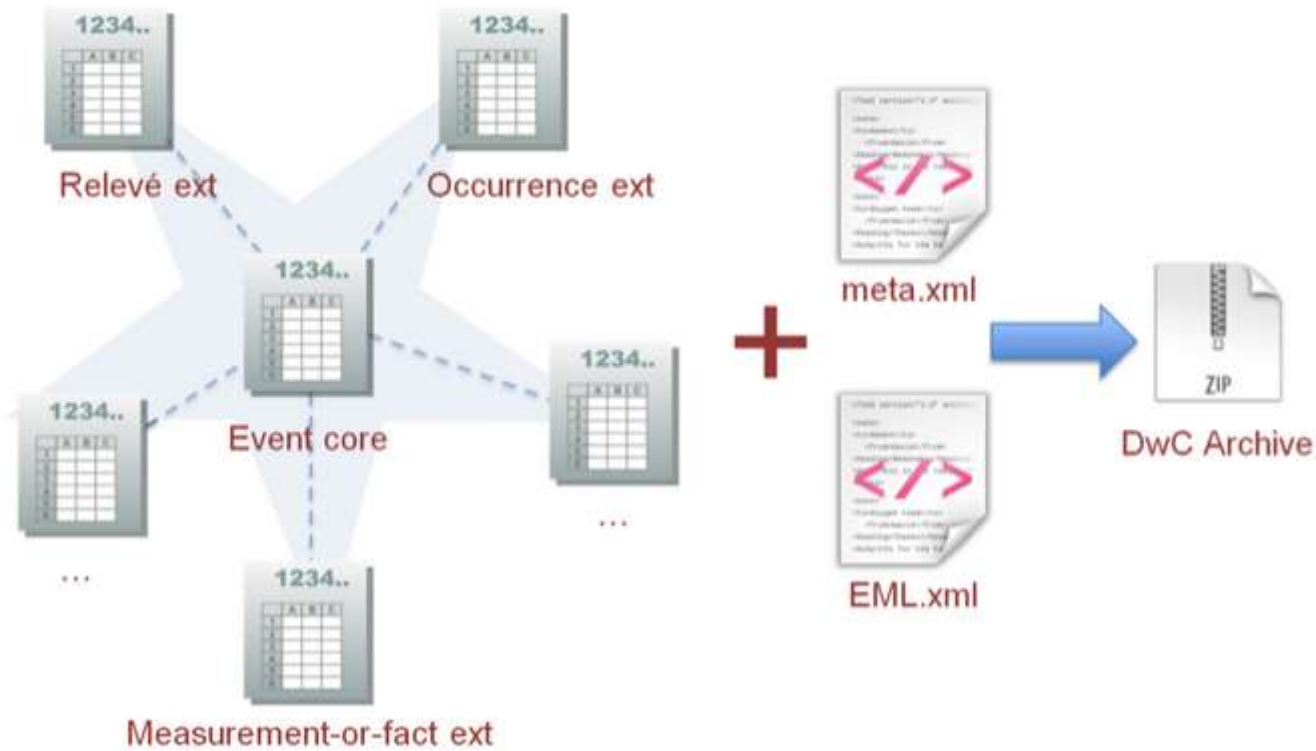
What is sampling-event data?

Sampling-event data describes species occurrences in time and space together with details of sampling effort. Such data is available from thousands of environmental, ecological, and natural resource investigations. These can be one-off studies or monitoring programmes. Such data are usually quantitative, calibrated, and follow certain protocols so that changes and trends of populations can be detected. This is in contrast to opportunistic observation and collection data, which today form a significant proportion of openly accessible biodiversity data.

<https://www.gbif.org/sampling-event-data>



Processing service to publish data products (digital objects) via GBIF Sampling-Event planned



SDD

EML

SDD-EML-Archive

<https://www.gbif.org/sampling-event-data>



Summarizing:

Several processing services to publish data products as digital objects are on the way...



NFDI 4
BIODIVERSITY
BIODIVERSITY, ECOLOGY & ENVIRONMENTAL DATA



Thank you!

