



# Datenmanagementpläne in DiversityDescriptions

Eine erste Analyse zur Service-Übergabe von  
GFBio an NFDI4Biodiversity

Tanja Weibulat  
IT Center  
Botanische Staatssammlung München

43. DWB-Workshop, 22.06.2022

SNSB   
Botanische  
Staatssammlung  
München

 **snsb  
it center**

# Datenmanagementpläne

- Zentral in Forschungsprojekten: Erhebung, Strukturierung, Analyse und Bewertung von Daten
- Selbstverständlich: Publikation erlangter Erkenntnisse in Textform (Publikationsoutput), nicht so der Daten
- DMP: Strukturiert Umgang mit Forschungsdaten, dokumentiert Verfahren mit den erhobenen Daten während und nach Abschluss eines Forschungsprojektes
- Forschungsförderer: Bedingung für die Bewilligung von Förderanträgen

Service von



NFDI 4 BIODIVERSITY



About Services Infothek Events GFBio e.V.

## GFBio Data Management Plan Tool

**BASIC QUESTIONS ABOUT DATA MANAGEMENT IN ONE PLACE**

**DYNAMIC DATA MANAGEMENT PLAN CREATION**

**PERSONAL SUPPORT**



Get a [free GFBio account](#) to save your dynamic DMP.

Please note: Requests that reach our #GFBioHelpdesk, #GFBioDMPT and #GFBioSubmission Service between 2021 delayed response due to reduced staff capacities over the holidays. We wish you all the best!

### Features and Services

#### Registered Users

- ✓ Answer important questions about the data management of your project
- ✓ Download your entries as a structured PDF file
- ✓ Request free personal [DMP support](#) from our experts
- ✓ Save your entries to your personal account and revise it whenever needed

#### Unregistered Users

- ✓ Answer important questions about your project
- ✓ Download your entries as a structured PDF file
- ✓ Request free personal [DMP support](#) from our experts

### Why write a DMP?

In short, a DMP helps you to ask all the right questions concerning data management. By writing a DMP, you save a lot of time in the long run, you are aware of potential data management obstacles, and you increase the transparency as well as the integrity of your work. A well-structured Data Management Plan (DMP) clarifies how and what data will be created, processed, and documented. It names means of data archiving and publication regarding costs as well as access conditions for the scientific community and the public. Also, DMPs are increasingly required as a mandatory proposal part by research institutions and funding agencies.

### Learn more about DMPT

In this how-to, you learn about (DMP) with the GFBio Data Management Plan Tool. The standard computer-produced DMP is a mandatory proposal part by research institutions and funding agencies.

### What does the DMPT create?

The standard computer-produced DMP is a mandatory proposal part by research institutions and funding agencies.



Welcome Tanja!

About Services Infothek Events GFBio e.V.

Services / Plan / DMPT /

- 1. General Project Information
- 2. Data Collection
- 3. Documentation and Metadata
- 4. Ethics and Legal Compliance
- 5. Preservation and Sharing

What is the official name of your research project? \*

Project Name

Please select a category:

Select

Is your research data reproducible? ●

One-time observation

Repeatable experiments

Time series

Add additional information (e.g. data reproduction might cause high costs or a lot of effort).

Please specify your project type. ●

Field Work

Simulation

Observational

Assimilation

Experimental

Modelling

Laboratory

Other

Provide your project abstract or describe your work and the data involved.

Project Abstract

Who is the point of contact for the project data? \* ●

Tanja Hofmann

Phone Number

hugo.dionex@arcor.de

Who are the principal investigators?

Same as point of contact for project data.

Principal Investigator

Add principal investigator

For which funding are you applying? ●

Select



Welcome Tanja!

About Services Infothek Events GFBio e.V.

Services / Plan / DMPT /

- 1. General Project Information
- 2. Data Collection
- 3. Documentation and Metadata
- 4. Ethics and Legal Compliance
- 5. Preservation and Sharing

Do you want to submit physical objects along with your data?

Yes

No

Is your object dead or alive?

Alive

Dead

Is your object taxon-based?

Yes

No

Do you have mainly sequence data?

Yes

No

What type of data will you create? ●

Text (notes, surveys, etc.)

Numeric (spreadsheet, measurements, etc.)

Models, code

Multimedia (images, sounds, video, etc.)

GIS data

Molecular Sequence Data

Other

What file formats will you create? ●

List all data formats you are planning to produce and give additional information if necessary (such as "only legible with software xy"). There are loads of data formats and some disciplines even have their own ones.

Please estimate the data volume you will create. ●

Cannot estimate

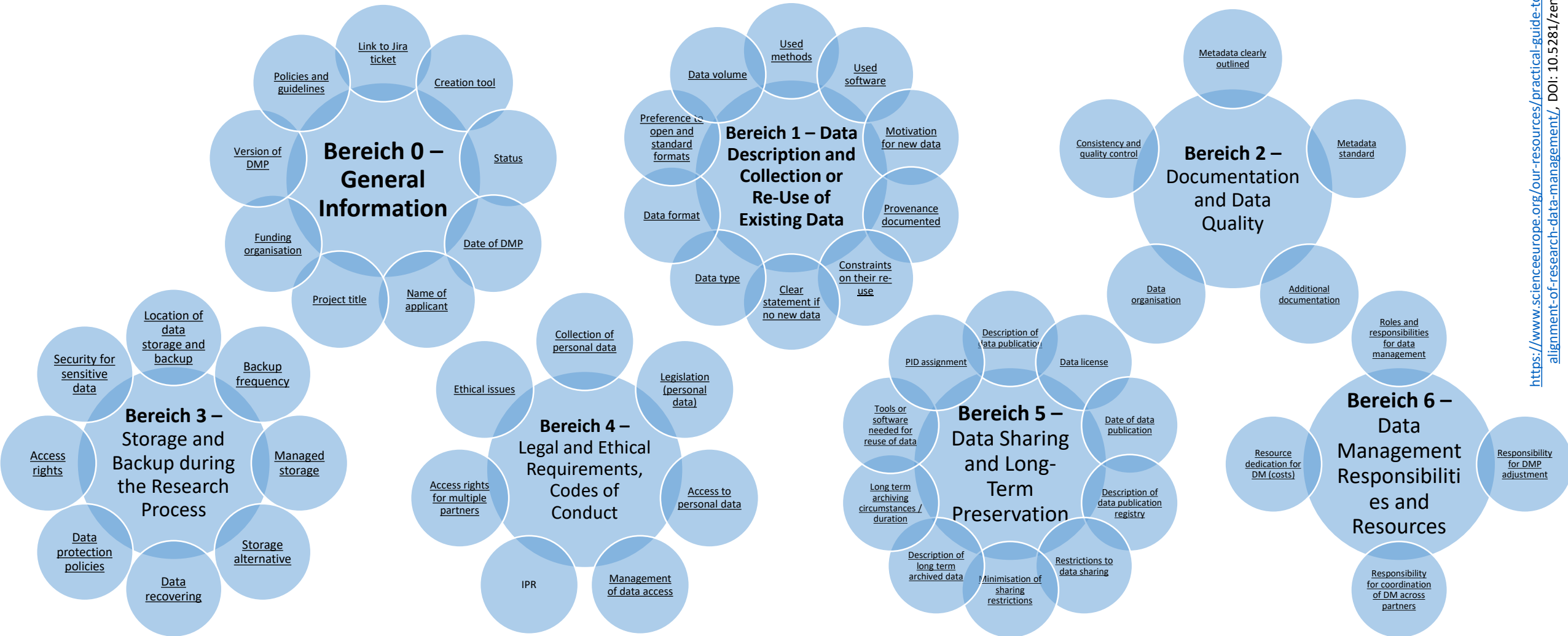
Please estimate the number of data sets (files) you will create. ●

Cannot estimate

Tanja Weibulat, 43. DWB-Workshop, 22.06.2022



# Evaluationskriterien



# Evaluationskriterien in DD

The screenshot displays the DiversityDescriptions software interface, version 4.5.3, with the 'Project title' descriptor configuration window open. The interface includes a menu bar (Connection, Edit, Grid, Query, Data, Administration, Help) and a toolbar. On the left, a 'Query results' pane shows a list of 25 evaluation criteria, with '6 - Project title' selected. Below this is a 'Query conditions' section for the 'Descriptor'.

The main configuration area for the 'Project title' descriptor includes the following fields and options:

- Descriptor:** Project title
- No:** 6
- ID:** 122897
- Main attributes:** Categorical states, Descriptor tree, Modifier/Frequency, Resource links
- Abbrev.:** (empty text field)
- Details:** Angabe des Projekttitels. Bei Science Europe wird die Projektnummer als wünschenswert angegeben, diese entfällt bei den GFBio-DMPs jedoch, da sie zur Antragstellung erstellt wurden. Eine Projektnummer wäre erst zu einem späteren Zeitpunkt nach Genehmigung eintragbar, solche späteren DMPs wurden aber gegebenenfalls von den Wissenschaftler\*innen selbst erstellt und sind nicht über GFBio erstellt worden.
- Notes:** (empty text field)
- Natural language support:** Wording: (empty text field)
- before:** (empty text field)
- after:** (empty text field)
- Descriptor type:**  Categorical,  Quantitative,  Text,  Sequence
- Descriptor settings:** Statistical measurement scale: Nominal, State collection model: OrSet, Repeatability: Standard, Availability: Standard
- Measurement data:** Prefix: (empty), Base: (empty), Unit: (empty), Unit is prefix:  Minimum: -1,790000E+308, Maximum: 1,790000E+308, Values are integer:
- Molecular data:**  Nucleotide sequence,  Protein sequence, Enable ambiguity:  Symbol length: 1, Gap symbol: (empty)
- Other options:**  Exclusive,  Mandatory

At the bottom left, the 'Query conditions' section shows the following configuration:

- Descriptor:** Title ~ (empty), Details ~ (empty), Project = GFBioDMPdesc

# Evaluationskriterien in DD

The screenshot displays the DiversityDescriptions software interface, version 4.5.3, with the 'Data license' descriptor selected. The interface is divided into several panels:

- Left Panel:** A list of 53 evaluation criteria. '41 - Data license' is highlighted in blue.
- Top Panel:** Shows the descriptor name 'Data license', its number '41', and its ID '122972'.
- Table:** A table with 5 columns: 'No.', 'Categorical state', 'Abbreviation', 'Details', and 'Wording'. It lists 12 different license types.
- Bottom Panel:** Includes 'State resource links' and 'Resources' sections with input fields for URL, Type, and Variant.

No.	Categorical state	Abbreviation	Details	Wording
1	CC BY		Creative Commons Attribution-4.0	
2	CC BY-NC		Creative Commons Attribution-NonCommercial-4.0	
3	CC BY-NC-ND		Creative Commons Attribution-NonCommercial-No...	
4	CC BY-NC-SA		Creative Commons Attribution-NonCommercial-Sha...	
5	CC BY-ND		Creative Commons Attribution-NoDerivs-4.0	
6	CC BY-SA		Creative Commons Attribution-ShareAlike-4.0	
7	CC0		Creative Commons Zero-1.0	
8	CC (not further specified)			
9	Other License			
10	License not further specified			
11	No license			
12	Not specified			

# Evaluationskriterien in DD

The screenshot displays the DiversityDescriptions software interface. The main window shows the 'Data license' descriptor, with a table listing various license types. The 'Descriptor tree' on the right side of the interface provides a hierarchical view of the evaluation criteria, including categories like 'General Information', 'Data Description and Collection or Re-Use of Existing Data', 'Documentation and Data Quality', 'Storage and Backup during the Research Process', 'Legal and Ethical Requirements, Codes of Conduct', 'Data Sharing and Long-Term Preservation', and 'Data Management Responsibilities and Resources'.

No.	Categorical state	Abbreviation	Details	Wording
1	CC BY		Creative Commons Attribution-4.0	
2	CC BY-NC		Creative Commons Attribution-NonCommercial-4.0	
3	CC BY-NC-ND		Creative Commons Attribution-NonCommercial-No...	
4	CC BY-NC-SA		Creative Commons Attribution-NonCommercial-Sha...	
5	CC BY-ND		Creative Commons Attribution-NoDerivs-4.0	
6	CC BY-SA		Creative Commons Attribution-ShareAlike-4.0	
7	CC0		Creative Commons Zero-1.0	
8	CC (not further specified)			
9	Other License			
10	License not further specified			
11	No license			
12	Not specified			

# Ergebnistabelle in DD

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Description title	Link to JIRA ticket	Creation tool	Status	Date of DMP	Name of applicant	Project title	Funding organisation	Version of DMP	Policies and guidelines	Used methods	Used software	Motivation for new data	Provenance documented	Constraints on their re-use	Clear statement if no new data	Data type	Data format	Preference to open and standard formats	Data volume [GB]	Metadata clearly outline
GFBio_DMP_00306	<a href="https://helpdesk.gfbio.org/browse/HELP-306">https://helpdesk.gfbio.org/browse/HELP-306</a>	1	1	13.02.2018	1	1	3	2	1	1	3	2	3	3	3	4/6/7/9	30	1	100	2
GFBio_DMP_00422	<a href="https://helpdesk.gfbio.org/browse/HELP-422">https://helpdesk.gfbio.org/browse/HELP-422</a>	2	1	30.08.2018	1	1	3	2	1	2	3	2	3	3	3	6	30	3	12000	2
GFBio_DMP_00436	<a href="https://helpdesk.gfbio.org/browse/HELP-436">https://helpdesk.gfbio.org/browse/HELP-436</a>	2	1	31.08.2018	1	1	3	2	1	1	3	2	3	3	3	4	22	2	1000	2
GFBio_DMP_00445	<a href="https://helpdesk.gfbio.org/browse/HELP-445">https://helpdesk.gfbio.org/browse/HELP-445</a>	2	1	07.09.2018	1	1	3	2	1	2	3	2	3	3	3	3/6	30	3	1000	2
GFBio_DMP_00450	<a href="https://helpdesk.gfbio.org/browse/HELP-450">https://helpdesk.gfbio.org/browse/HELP-450</a>	2	1	13.09.2018	1	1	3	2	1	1	3	2	3	3	3	3/6/8/9/13	2/24	2	10	2
GFBio_DMP_00537	<a href="https://helpdesk.gfbio.org/browse/HELP-537">https://helpdesk.gfbio.org/browse/HELP-537</a>	2	1	20.12.2018	1	1	3	2	1	1	1	2	3	3	3	3	2/4/5/10/24/26/28/29	2	1000	1
GFBio_DMP_00802	<a href="https://helpdesk.gfbio.org/browse/HELP-802">https://helpdesk.gfbio.org/browse/HELP-802</a>	2	1	14.09.2020	1	1	3	2	1	1	1	2	3	3	3	2/3/4/5/12/14	2/22/28	2	500	1
GFBio_DMP_00811	<a href="https://helpdesk.gfbio.org/browse/HELP-811">https://helpdesk.gfbio.org/browse/HELP-811</a>	2	1	14.09.2020	1	1	3	2	1	1	2	2	3	3	3	4/6/9/12/14	1/2/9/15/16/21/22/24/28	3	4500	1
GFBio_DMP_00876	<a href="https://helpdesk.gfbio.org/browse/HELP-876">https://helpdesk.gfbio.org/browse/HELP-876</a>	2	1	12.04.2019	1	1	3	2	1	1	1	2	3	3	3	4/6/12	2/5/7/28	3	21	1
GFBio_DMP_00988	<a href="https://helpdesk.gfbio.org/browse/HELP-988">https://helpdesk.gfbio.org/browse/HELP-988</a>	2	1	27.01.2020	1	1	3	2	1	1	1	2	3	3	3	3/6/12/14/16	1/2/6/15/16/24/29	3	100	1
GFBio_DMP_01000	<a href="https://helpdesk.gfbio.org/browse/HELP-1000">https://helpdesk.gfbio.org/browse/HELP-1000</a>	2	1	27.01.2020	1	1	3	2	1	1	1	2	3	3	3	6/9/14	2/7/8/9/17/25/28/29	3	1000	2
GFBio_DMP_01004	<a href="https://helpdesk.gfbio.org/browse/HELP-1004">https://helpdesk.gfbio.org/browse/HELP-1004</a>	2	1	29.01.2020	1	1	3	2	1	1	1	2	2	2	3	5/6/9/13	2/6/24/29	3	1000	1
GFBio_DMP_01009	<a href="https://helpdesk.gfbio.org/browse/HELP-1009">https://helpdesk.gfbio.org/browse/HELP-1009</a>	2	1	27.01.2020	1	1	3	2	1	1	1	1	1	2	3	4/5/6/9/12/16	2/6/7/9/13/14/24/29	2	2000	1
GFBio_DMP_01010	<a href="https://helpdesk.gfbio.org/browse/HELP-1010">https://helpdesk.gfbio.org/browse/HELP-1010</a>	2	1	27.01.2020	1	1	3	2	1	1	1	2	2	2	3	1/6/9/12/14/16	2/6/7/23/27/29	3	10000	1
GFBio_DMP_01011	<a href="https://helpdesk.gfbio.org/browse/HELP-1011">https://helpdesk.gfbio.org/browse/HELP-1011</a>	2	1	27.01.2020	1	1	3	2	1	1	2	2	3	3	3	6/9/12/14	2/7/13/14/28/29	3	1000	2
GFBio_DMP_01032	<a href="https://helpdesk.gfbio.org/browse/HELP-1032">https://helpdesk.gfbio.org/browse/HELP-1032</a>	2	1	25.02.2020	1	1	3	2	1	1	1	2	3	3	3	4/6/9/12/14/16	2/6/9/15/28/29	2	1	1
GFBio_DMP_01034	<a href="https://helpdesk.gfbio.org/browse/HELP-1034">https://helpdesk.gfbio.org/browse/HELP-1034</a>	2	1	07.04.2020	1	1	4	2	1	1	1	2	3	3	3	3/4/5/16	2/4/9/24/28/29	2	24	1
GFBio_DMP_01089	<a href="https://helpdesk.gfbio.org/browse/HELP-1089">https://helpdesk.gfbio.org/browse/HELP-1089</a>	2	1	15.07.2020	1	1	3	2	1	1	1	1	3	3	3	4/6	1/2/7/9/15/16/22/24	1	162	1
GFBio_DMP_01115	<a href="https://helpdesk.gfbio.org/browse/HELP-1115">https://helpdesk.gfbio.org/browse/HELP-1115</a>	2	1	08.09.2020	1	1	3	2	1	1	1	2	3	3	3	3/6/9/12/16	3/7/24/28	3	95000	2
GFBio_DMP_01124	<a href="https://helpdesk.gfbio.org/browse/HELP-1124">https://helpdesk.gfbio.org/browse/HELP-1124</a>	2	1	05.02.2021	1	1	2	2	1	1	1	3	3	3	3	5/7/9/10/11/16	2/5/9/11/12/22/28/29	2	1000	1
GFBio_DMP_01125	<a href="https://helpdesk.gfbio.org/browse/HELP-1125">https://helpdesk.gfbio.org/browse/HELP-1125</a>	2	1	22.09.2020	1	1	5	2	1	1	3	1	1	2	3	3/4	4/24/28	2	1,5	1
GFBio_DMP_01143	<a href="https://helpdesk.gfbio.org/browse/HELP-1143">https://helpdesk.gfbio.org/browse/HELP-1143</a>	2	1	09.12.2020	1	1	3	2	1	1	2	2	3	3	3	4/12/14	2/5/9/20/24/28/29	2	9	1
GFBio_DMP_01159	<a href="https://helpdesk.gfbio.org/browse/HELP-1159">https://helpdesk.gfbio.org/browse/HELP-1159</a>	2	1	07.10.2021	1	1	3	2	1	1	2	2	3	3	3	3/4/5/6/9/12/14/16	2/5/6/7/20/22/24/29	3	10	1
GFBio_DMP_01203	<a href="https://helpdesk.gfbio.org/browse/HELP-1203">https://helpdesk.gfbio.org/browse/HELP-1203</a>	2	1	20.07.2021	1	1	3	2	1	1	2	2	3	3	3	3/4/6/12/13/16	2/4/5/9/15/18/22/24/26/28/29	3	10000	1
GFBio_DMP_01204	<a href="https://helpdesk.gfbio.org/browse/HELP-1204">https://helpdesk.gfbio.org/browse/HELP-1204</a>	2	1	02.03.2021	1	1	3	2	1	1	3	2	3	3	3	3	2/4/5/10/24/26/28/29	2	7	1



# Auswahl – Ergebnisse

- Verteilung über die Jahre
  - TaxonOmics?
  - Lockdown?
  - NFDI?
- Häufigkeit der meistgenutzten Dateiformate
  - 29 verschiedene Dateiformate.
  - Nicht für alle Datentypen in einem DMP angegeben, aber in jedem DMP Angaben vorhanden.
  - Am häufigsten verwendete Dateiformate sind die am weitesten verbreiteten, nicht spezialisierte.

Jahr der Herausgabe	Anzahl der DMPs
2018	6
2019	1
2020	14
2021	4

