

help text in English - Hackathon consolidated Version

Text:

The following questions deal with the content, format, and origin of the research data generated or reused in the project and describe their management. The best way to address these questions is by clustering the data in logical units. For each of those "data sets" the same questions will be asked. Data sets may represent rather broader categories with similar constraints/properties regarding data management.

Assign a meaningful but short name to each "data set" with reference to the research question, and use it to label the corresponding tab in RDMO.

%%%

In an RDMO "akkordeon"

%%%

You can proceed like that:

- Gain an overview of the data being generated or used within your research activity.
- Evaluate each data category/set with respect to the following criteria:
 - origin/generation: generated (observation/measurement, experiment, survey, simulation, web-crawl,...) or reused;
 - storage (space) requirements: e.g.: <1GB, ~10TB
 - documentation requirements: exampleA: extensive description of survey sampling and survey implementation; exampleB: reference/link to existing meta-data to re-used data retrieved from a research data repository
 - data privacy aspects: personal data; sensitive non-personal; public data;
 - relevance / reuse potential: high (reference data), middle (scientific main study), low (preliminary study)
 - publication
 - long term preservation

Whenever two data sets substantially differ from each other considering these criteria, they should be regarded as different categories.

Examples of two project with data organised in categories is to be found on this [GitLab page](https://gl.mathhub.info/cdi/labs/literacy/proposal-self-service/-/blob/main/RDMO/Beispiele_un_d_Hilfen/Hilfe_Datenkategorie_en.md) [English]

Mockups:

Data Sets

The following questions deal with the content, format, and origin of the research data generated or reused in the project and describe their management. The best way to address these questions is by clustering the data in logical units. For each of those "data sets" the same questions will be asked. Data sets may represent rather broader categories with similar constraints/properties regarding data management.

Assign a meaningful but short name to each "data set" with reference to the research question, and use it to label the corresponding tab in RDMO.

► Some tips for choosing data sets

Please fill in the form for each tab. The same tabs may be used later on other pages. You can add a new tab using the green button. Once created, you can edit or delete tabs using the buttons in the top right corner.

Data Set I

+ Data set



Please give some keywords describing the data set.

Data Sets

The following questions deal with the content, format, and origin of the research data generated or reused in the project and describe their management. The best way to address these questions is by clustering the data in logical units. For each of those "data sets" the same questions will be asked. Data sets may represent rather broader categories with similar constraints/properties regarding data management.

Assign a meaningful but short name to each "data set" with reference to the research question, and use it to label the corresponding tab in RDMO.

▼ Some tips for choosing data sets

You can proceed like that: Gain an overview of the data being generated or used within your research activity. Evaluate each data category with respect to the following criteria:

- origin/generation: generated (observation/measurement, experiment, survey, simulation, web-crawl,...) or reused;
- storage (space) requirements: e.g.: <1GB, ~10TB
- documentation requirements: e.g. extensive description of survey sampling and survey implementation or detailed references to meta-data and re-used data retrieved from a research data repository
- data privacy aspects: personal data, sensitive non-personal, public data
- relevance / reuse potential: high (reference data), middle (scientific main study), low (preliminary study)
- data publication
- long term preservation

Examples of two project with data organised in categories is to be found [on the GitLab page](#).

Please fill in the form for each tab. The same tabs may be used later on other pages. You can add a new tab using the green button. Once created, you can edit or delete tabs using the buttons in the top right corner.

RDMO-HTML:

The following questions deal with the content, format, and origin of the research data generated or reused in the project and describe their management. The best way to address these questions is by clustering the data in logical units. For each of those "data sets" the same questions will be asked. Data sets may represent rather broader categories with similar constraints/properties regarding data management.

Assign a meaningful but short name to each "data set" with reference to the research question, and use it to label the corresponding tab in RDMO.

<details>

<summary style="cursor:pointer">Some tips for choosing data sets</summary>

<p>You can proceed like that: Gain an overview of the data being generated or used within your research activity.

Evaluate each data category with respect to the following criteria:

-
 - origin/generation: generated (observation/measurement, experiment, survey, simulation, web-crawl,...) or reused;
 - storage (space) requirements: e.g.: <1GB, ~10TB
 - documentation requirements: e.g. extensive description of survey sampling and survey implementation or detailed references to meta-data and re-used data retrieved from a research data repository
 - data privacy aspects: personal data, sensitive non-personal, public data
 - relevance / reuse potential: high (reference data), middle (scientific main study), low (preliminary study)
 - data publication
 - long term preservation

</p>

<p>

Examples of two project with data organised in categories is to be found

on the GitLab page

</p>

</details>

Achtung: Es gibt in RDMO wohl etwas besseres als <details> </details> ; Mit {% ??? %} als Begrenzung, aber ich erinnere mich nicht. Jürgen Windeck weiß es aber.

help text in Deutsch- Hackathon consolidated Version

Die folgenden Fragen befassen sich mit dem Inhalt, dem Format und der Herkunft der im Projekt generierten oder wiederverwendeten Forschungsdaten und beschreiben deren Verwaltung. Am besten lassen sich diese Fragen beantworten, indem die Daten in logischen Einheiten zusammengefasst werden. Für jeden dieser „Datensätze“ werden dieselben Fragen gestellt. Datensätze können eher breitere Kategorien mit ähnlichen Einschränkungen/Eigenschaften hinsichtlich der Datenverwaltung darstellen.

Weisen Sie jedem „Datensatz“ einen aussagekräftigen, aber kurzen Namen zu, der sich auf die Forschungsfrage bezieht, und verwenden Sie diesen Namen, um den entsprechenden Tab in RDMO zu beschriften.

%%%

In “akkordeon”

%%%

Einige Tipps zur Auswahl von Datensätzen

Sie können wie folgt vorgehen: Verschaffen Sie sich einen Überblick über die Daten, die im Rahmen Ihrer Forschungsaktivitäten generiert oder verwendet werden. Bewerten Sie jede Datenkategorie anhand der folgenden Kriterien:

- Herkunft/Generierung: generiert (Beobachtung/Messung, Experiment, Umfrage, Simulation, Web-Crawl,...) oder wiederverwendet;
- Speicherplatzanforderungen: z. B.: <1 GB, ~10 TB
- Dokumentationsanforderungen: z. B. ausführliche Beschreibung der Stichprobenziehung und Durchführung der Umfrage oder detaillierte Verweise auf Metadaten und wiederverwendete Daten, die aus einem Forschungsdaten-Archiv abgerufen wurden
- Datenschutzaspekte: personenbezogene Daten, sensible nicht personenbezogene Daten, öffentliche Daten
- Relevanz/Wiederverwendungspotenzial: hoch (Referenzdaten), mittel (wissenschaftliche Hauptstudie), gering (Vorstudie)
- Veröffentlichung der Daten
- Langzeitarchivierung

Beispiele für zwei Projekte mit nach Kategorien geordneten Daten finden Sie auf der [GitLab-Seite](

https://gl.mathhub.info/cdi/labs/literacy/proposal-self-service/-/blob/main/RDMO/Beispiele_und_Hilfen/Hilfe_Datenkategorie_de.md)

Version with hackathon edits

help text in English

The following questions deal with the content, format, and origin of the research data generated or reused in the project and describe their management. The best way to address these questions is by clustering the data in a small number of homogeneous "data categories".

You can proceed like that:

- Gain an overview of the data being generated or used within your research activity.
 - Evaluate each data category/set with respect to the following criteria:
 - origin/generation: generated (observation/measurement, experiment, survey, simulation, web-crawl,...) or reused;
 - storage (space) requirements: e.g.: <1GB, ~10TB
 - documentation requirements: e.g. or ... or ...
 - data privacy aspects: personal data, sensitive non-personal, public data,
 - relevance / reuse potential: high (reference data), middle (scientific main study), low (preliminary study)
- Whenever two data sets substantially differ from each other considering these criteria, they should be regarded as different categories.

- Assign a meaningful but short name to each "data category" with reference to the research question, and use it to label the corresponding tab in RDMO. Then describe each data category using the checkboxes/text fields on this page and the following ones.

Please fill in the form for each tab. The same tabs may be used later on other pages.

Once created, you can edit or delete tabs using the buttons in the top right corner.

Datensatz01

+ Set

Example of effective category names:

e.g. quantitative survey data; qualitative interview data; choice experiment data; simulation data; farm survey data; production data; climate data etc.;

Examples of two project with data organised in categories is to be found here:

https://gl.mathhub.info/cdi/labs/literacy/proposal-self-service/-/blob/main/RDMO/Beispiele_un_d_Hilfen/Hilfe_Datenkategorie_en.md [English]

https://gl.mathhub.info/cdi/labs/literacy/proposal-self-service/-/blob/main/RDMO/Beispiele_un_d_Hilfen/Hilfe_Datenkategorie_de.md [German]

Please fill in the form for each tab. The same tabs may be used later on other pages. You can add a new tab using the green button. Once created, you can edit or delete tabs using the buttons in the top right corner.

Mockup:

Data Category

The following questions deal with the content, format, and origin of the research data generated or reused in the project and describe their management. The best way to address these questions is by clustering the data in a small number of homogeneous "data categories".

► Some tips for choosing data categories

Please fill in the form for each tab. The same tabs may be used later on other pages. You can add a new tab using the green button. Once created, you can edit or delete tabs using the buttons in the top right corner.

Datenkategorie I

+ Data category



Please give some keywords describing the data set.

Please enter your entries line by line. You can add lines using the green button and remove them using the blue cross (x).

+ Keyword

Data Category

The following questions deal with the content, format, and origin of the research data generated or reused in the project and describe their management. The best way to address these questions is by clustering the data in a small number of homogeneous "data categories".

▼ Some tips for choosing data categories

You can proceed like that: Gain an overview of the data being generated or used within your research activity. Evaluate each data category with respect to the following criteria:

- origin/generation: generated (observation/measurement, experiment, survey, simulation, web-crawl,...) or reused;
- storage (space) requirements: e.g.: <1GB, ~10TB
- documentation requirements: e.g. extensive description of survey sampling and survey implementation or ... reference to meta-data to re-used data retrieved from a research data repository or ...
- data privacy aspects: personal data, sensitive non-personal, public data
- relevance / reuse potential: high (reference data - we intend to publish the data), middle (scientific main study - we have no plans to publish, yet), low (preliminary study - we do not be published)

Whenever two data sets substantially differ from each other considering these criteria, they should be regarded as different categories. Also, assign a meaningful but short name to each "data category" with reference to the research question, and use it to label the corresponding tab in RDMO. Then describe each data category using the checkboxes/text fields on this page and the following ones

Two examples:

- A medical research project decides to write an overarching DMP for all subprojects. This DMP only give the broad plans of the project on data storage, dissemination and data protection. PhD students are expected to provide supporting DMP for their project that will address details like preferred data formats, data cleansing and processing. The "global DMP" therefore uses very broad categories of data. Here is a link with more details on the data and the reasoning: <https://pad.gwdg.de/BbuoxtyfTAGxrsCF-9Q2bw?both#>.
- A project studies communication and regulations during the COVID pandemic. The DMP is written as part of a grant application. Legal aspects (data protection, copyright, ...) will be of high importance - more so since the evaluation panel will involve legal scholars. Therefore legal aspects (and file size) become the key discrimination factors for deciding on the number of data categories. Here is a link with details on the data and the reasoning: <https://pad.gwdg.de/BbuoxtyfTAGxrsCF-9Q2bw?both#>.

Please fill in the form for each tab. The same tabs may be used later on other pages. You can add a new tab using the green button. Once created, you can edit or delete tabs using the buttons in the top right corner.

Datenkategorie I

+ Data category



Please give some keywords describing the data set.

Hilfetext auf Deutsch

.....

Neuer Namen und Implementierung in RDMO

Bevorzugte Benennung: "Datenkategorie" / "Data category" ?

Wo es geändert wird:

Das Etikett zum Reiter ist im jeweiligen Katalog enthalten, bspw.

<https://github.com/rdmorganiser/rdmo-catalog/blob/main/shared/metrology-rdm/epm-que>

[stions.xml#L292](#) , kann dann geändert werden und wird durch "Set" ersetzt, falls nicht gegeben.

Wo es nicht geändert werden kann:

Das Etikett zum Knopf muss in der Software hartkodiert sein und ist im Software-Repo lokalisiert:

https://github.com/rdmorganiser/rdmo/blob/main/rdmo/locale/de/LC_MESSAGES/django.po#L2447 Es wurde absichtlich mit dem allgemeinen Wort "Entry/Eintrag" belegt und kann nicht verändert werden.

#1 + Dataset group

Please list all datasets belonging to this group.

Please enter your entries line by line. You can add lines using the green button and remove them using the blue cross (x).

+ Entry

#1 + Set

Bitte nutzen Sie für die Einträge jeweils eine eigene Zeile. Sie können weitere Einträge mit dem grünen Button hinzufügen und mit dem blauen Kreuz (x) entfernen.

+ Eintrag