Catalog Horizon2020

Comment Bei diesem Fragebogen wurde eine neue Sortierung erarbeitet, die einen weiteren Zugang zu den Fragen für ein Horizon2020-DMP bieten soll. Während das Template von Horizon2020 die FAIR data Prinzipien im Vordergrund hat, soll die Sortierung in diesem Fragebogen inhaltliche Zusammenhänge darstellen und so die Beantwortung der Fragen zu erleichtern. Bei der jeweiligen Frage befindet sich im Kommentarfeld ein Hinweis auf welche Original-H2020-Frage die aktuelle Frage in RDMO abzielt. Die Nummerierung existiert so im H2020-Template nicht, sie wurde zum besseren Überblick ergänzt.

Section 1. General information about the research project

Question set Purpose of the project

Questions:

- **Text:** 1.1 What is the purpose of the data collection/generation and its relation to the objectives of the project?

  Help: Please give an overview of the aim of the project and describe the individual work packages or sub-projects in which data are collected.

  Comment: H2020: 1.1 What is the purpose of the data collection/generation and its relation to the objectives of the project?

Question set Contact person

Questions:

- **Text:** 1.2 Who will be responsible for data management in your project?

  Help: Please give the name and an email address.

  Comment: H2020: 3.3 Who will be responsible for data management in your project?

Question set Processing status

Questions:

- **Text:** 1.3 Versionnumber and date

  Help: Please note version number and date of last development (e.g. V01_2018-12-12)

  Comment: not included in Original-H2020-Template

Section 2. Existing data

Question set Data

Questions:

- **Text:** 2.1. Is it planned to integrate or reuse existing data?

  Help: If existing data are reused, please describe under which conditions they can be used (which institution provides the data ; which requirements are to take care of e.g. proposals, agreements for data reuse).

  Comment: H2020: 1.3 Will you re-use any existing data and how?

- **Text:** 2.2 When data is reused, who created the record?

  Help: Please specify under which address, PID or URL the dataset is already available.

Comment: H2020: 1.4 What is the origin of the data?

- **Text:** 2.3 Are the data used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers "DOI")? If yes, which meachnism is used?

  Help: Please describe them:

  Comment: H2020: 2.1.1 (for used data) Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)

Section 3. Generated data

Question set Data

Questions:

- **Text:** 3.1 What types and formats of data will the project generate/collect?

  Help: When choosing a data format, one should consider the consequences for collaborative use, long-term preservation as well as re-use. It is advisable to prefer formats that are standardised, open, non-proprietary and well-established in the respective scholarly community. More criteria and detailed explanations can be found e.g. in the WissGrid-Leitfaden, pp. 22 f.).

  Comment: H2020: 1.2 What types and formats of data will the project generate/collect?

- **Text:** 3.2 What is the actual or expected size of the dataset?

  Help: Please estimate the size of the dataset (it is important for the storage during or longterm preservation after the project).

  Comment: H2020: 1.5 What is the expected size of the data?

Section 4. Metadata and referencing

Question set Metadata

Questions:

- **Text:** 4.1 What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

  Comment: H2020: 2.1.5 What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

- **Text:** 4.1.1 Which metadata will be generated automatically?

  Comment: H2020: 2.1.5 What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

- **Text:** 4.1.2 Which metadata will be generated semi-automatically?

  Comment: H2020: 2.1.5 What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

- **Text:** 4.1.3 Which metadata will be created manually?

Comment: H2020: 2.1.5 What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

Question set Metadata standards

Questions:

- **Text:** 4.2 What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?

  Comment: H2020: 2.3.2 What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?

- **Text:** 4.3 Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?

  Comment: H2020: 2.3.3 Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?

- **Text:** 4.4 In case it is unavoidable that you use uncommon or generate project-specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?

  Comment: H2020: 2.3.4 In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?

Question set PIDs

Questions:

- **Text:** 4.5 Are the data produced in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers "DOI")? If yes, which meachnism is used?

  Help: Please describe:

  Comment: H2020: 2.1.1 (generated data) Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?

Question set Subject heading

Questions:

- **Text:** 4.6 Will search keywords be provided that optimize possibilities for re-use of data?

  Help: These keywords are intended to make it possible to find the data for subsequent use.

  Comment: H2020: 2.1.3 Will search keywords be provided that optimize possibilities for re-use?

Section 5. Data organisation and security

Question set Guidelines

Questions:

- **Text:** 5.1 Is there an internal project policy for naming the data? If so, please outlinenaming conventions briefly and link to more detailed documentation if necessary.

  Help: Here you should describe how the data/files used in the project are named and how it is stored in a file/folder structure. It is recommended to define uniform rules for file storage,

directory and file naming as well as versioning in the project and to record these in the data management plan. This is important to ensure that the data used remains traceable during the research process and to avoid data loss or changes due to accidental deletion, overwriting or relocation. In addition, uniform naming reduces the effort required for later data archiving.

Comment: H2020: 2.1.2 What naming conventions do you follow?

Question set Version strategy

Questions:

- **Text:** 5.2 Which versioning strategy is used for data records?

  Help: Please describe briefly the internal project rules for versioning data sets (e.g: Which changes require a new version? How are the changes documented? How are the different versions named?)

  Comment: H2020: 2.1.4 Do you provide clear version numbers?

- **Text:** 5.3. Are there well described conditions for access (rights management)?

  Help: Please describe the procedures to protect the data against unallowed access (e.g. access authorisation and restrictions, encryption, password protection, firewall)

  Comment: H2020: 2.2.11 Are there well described conditions for access (i.e. a machine readable license)?

Question set Data security

Questions:

- **Text:** 5.4 If there are restrictions on use, how will access be provided?

  Comment: H2020: 2.2.9 If there are restrictions on use, how will access be provided?

- **Text:** 5.5 How will the identity of the person accessing the data be ascertained?

  Comment: H2020: 2.2.12 How will the identity of the person accessing the data be ascertained?

- **Text:** 5.6 What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?

  Comment: H2020: 4.1 What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?

Section 6. Legal and ethics

Question set Sensitive data

Questions:

- **Text:** 6.1 Are there any ethical or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).

  Help: It's to check if there are legal determinations e.g. data protection, rights of disposal or use. Legal arrangements influence in which way data can be generated, edited or stored. Please mention if there are rights of disposal and if there are personal data which underlie the protection of data privacy. There may also be copyrights on the used measuring devices.

Comment: H2020: 5.1 Are there any ethical or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).

Question set Personal data

Questions:

- **Text:** 6.2 Does the datasets contain personal data?

  Help: European Data Protection Law may apply. The German Federal Data Protection Act defines personal data as „any information concerning the personal or material circumstances of an identified or identifiable individual (the data subject)" (BDSG, §3 Abs. 1)). A person is "identified" when it is obvious whom the data is associated with. A person is "identifiable" when it is possible to find out which person the data refers to with additional information (See Häder 2009, p. 7). The handling of personal information is regulated by law. More information (in German only) can be found in the following publications: * Michael Häder (2009): Der Datenschutz in den Sozialwissenschaften. RatSWD Working Paper No. 90. * Uwe Jensen (2012): Leitlinien zum Management von Forschungsdaten. Sozialwissenschaftliche Umfragedaten. GESIS Technical Report 2012|07. (esp. pp. 13 ff) * Handreichung: Informationen zu rechtlichen Aspekten bei der Handhabung von Sprachkorpora. (esp. part 1.1 and part 2.2)

  Comment: H2020: 5.2 Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data?

Question set Acquiescence

Questions:

- **Text:** 6.3 Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data?

  Help: Basically, the collection, processing, archiving and publication of personal data is only admissible, when the "informed consent" of the person in question has been obtained. There are only very few cases in which this is not the case (see, also for more information: Informationen zu rechtlichen Aspekten bei der Handhabung von Sprachkorpora, p. 6).

  Comment: H2020: 5.2 Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data?

Section 7. Re-use

Question set Data access technology

Questions:

- **Text:** 7.1 What methods or software tools are needed to access and re-use the data?

  Help: To be able to re-use data (e.g. to replicate studies, for meta analysis or to solve new research questions), along with software, equipment and knowledge about special methods to use the data are required. Just as with the formats, the recommendation is: the more standardised, open and established, the better for re-use.

  Comment: H2020: 2.2.4 What methods or software tools are needed to access the data?

- **Text:** 7.2 Is documentation about the software needed to access the data included?

  Comment: H2020: 2.2.5 Is documentation about the software needed to access the data included?

- **Text:** 7.3 Is it possible to include the relevant software (e.g. in open source code)?

  Comment: H2020: 2.2.6 Is it possible to include the relevant software (e.g. in open source code)?

Question set Data access

Questions:

- **Text:** 7.4 For which persons, groups or institutions might the dataset be useful ('data utility')? What are possible scenarios?

  Comment: H2020: 1.6 To whom might it be useful ('data utility')?

Question set Embargo

Questions:

- **Text:** 7.5 When will the data be made available for re-use?

  Comment: H2020: 2.4.2 When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

- **Text:** 7.5.1 If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

  Comment: H2020: 2.4.3 Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.

Section 8. Storage and long-term preservation

Question set Interoperability

Comment wortlaut noch relativ nah an Original-H2020-Frage

Questions:

- **Text:** 8.1 Is the data set interoperable, i.e. suitable for data exchange and re-use between or by different researchers, institutions, organisations and countries?

  Help: (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)

Comment: H2020: 2.3.1 Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?

- **Text:** 8.2 Where will the data (including metadata, documentation, and any relevant code or software) be deposited? Preference should be given to certified repositories which support open access where possible.

  Help: Please indicate the procedures for data security and data privacy laws. This includes also information about storage location and backup. Preference should be given to certified repositories which support open access where possible.

  Comment: H2020: 2.2.3 How will the data be made accessible (e.g. by deposition in a repository)?

Question set Data publication and licencing

Questions:

- **Text:** 8.3 Which data produced and/or used in the project will be made openly available as the default? If certain data sets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.

  Help: Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if relevant provisions are made in the consortium agreement and are in line with the reasons for opting out.

  Comment: H2020: 2.2.1 Which data produced and/or used in the project will be made openly available as the default? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions. H2020: 2.2.2 Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if relevant provisions are made in the consortium agreement and are in line with the reasons for opting out.

- **Text:** 8.4 Under which conditions of use and which license should the data be published or shared to permit widest re-use possible?

  Comment: H2020: 2.4.1 How will the data be licensed to permit the widest re-use possible?

Question set Repository

Questions:

- **Text:** 8.5 Where is the data (including metadata, documentation, and any relevant code or software) stored or archived after the end of the project?

  Help: Will the metadata, the sourcecode and relevant software also be stored?

  Comment: H2020: 2.2.7 Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.

- **Text:** 8.6 Is the repository or data centre chosen certified (e.g. Data Seal of Approval, nestor Seal or ISO 16363)?

Help: If the dataset is archived at several places, you may answer this question with yes, if this applies to at least one of these.

Comment: H2020: 2.2.7 Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible. H2020: 4.2 Is the data safely stored in certified repositories for long term preservation and curation?

- **Text:** 8.7 Have you explored appropriate arrangements with the identified repository?

  Comment: H2020: 2.2.8 Have you explored appropriate arrangements with the identified repository?

Question set Data quality

Questions:

- **Text:** 8.8 Are data quality assurance processes described?

  Verbose name:

  - Verbose name: Data quality assurance
  - Verbose name plural:

  Comment: H2020: 2.4.5 Are data quality assurance processes described?

Question set Data access committee

Questions:

- **Text:** 8.9 Is there a need for a data access committee?

  Help: The data access committee decides if access to the project data is allowed.

  Comment: H2020: 2.2.10 Is there a need for a data access committee?

Question set Storage duration

Questions:

- **Text:** 8.10 How long is it intended that the data remains re-usable?

  Help: Please describe how long the data will be stored and what deadlines for deletion are in force.

  Comment: H2020: 2.4.4 How long is it intended that the data remains re-usable?

Section 9. Costs

Question set Additional support

Questions:

- **Text:** 9.1 Are there actual or potential usage scenarios that could benefit from support by a data management or IT expert, or that even require such support?

  Comment: H2020: 2.4.6 Further to the FAIR principles, DMPs should also address:

- **Text:** 9.2 Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?

Comment: H2020: 6.1 Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?

Question set FAIR - Costs

Questions:

- **Text:** 9.3 What are the personel costs for making data FAIR in your project?

  Comment: H2020: 3.1 What are the costs for making data FAIR in your project?

- **Text:** 9.4 What are the non-personnel costs for making data FAIR in your project?

  Help: If possible, please estimate the costs in **Euro**.

  Comment: H2020: 3.1 What are the costs for making data FAIR in your project?

Question set Costs for long term preservation

Questions:

- **Text:** 9.5 What are the personnel costs associated with long-term preservation for the project?

  Comment: H2020: 3.4 Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?

- **Text:** 9.6 What is the amount of non-personnel-costs regarding long-term preservation for the project?

  Help: If possible, please estimate the costs in **Euro**.

  Comment: H2020: 3.4 Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?

- **Text:** 9.7 Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?

  Comment: H2020: 3.4 Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?

Question set Cost

Questions:

- **Text:** 9.8 How will these costs be covered?

  Help: Note that costs related to open access to research data are eligible as part of the Horizon 2020 grant (if compliant with the Grant Agreement conditions).

  Comment: H2020: 3.2 How will these be covered? Note that costs related to open access to research data are eligible as part of the Horizon 2020 grant (if compliant with the Grant Agreement conditions).