UNIVERSITY OF COPENHAGEN

Annotated Horizon Europe DMP template for the University of Copenhagen (UCPH)

Version: 1.1
Date: 2023-10-05
Created by: Falco Hüser, Copenhagen University Library, falh@kb.dk
Comments: based on “Horizon Europe: Data Management Plan Template” (v1.1, 05 01.04.2022)
https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/temp-form/report/data-management-plan_he_en.docx

The template and all guidance are also available in the DeiC DMP tool for UCPH users: https://dmp.deic.dk/


This work is licensed under a Creative Commons Attribution 4.0 International License.

Horizon Europe
The Horizon Europe Model Grant Agreement requires that a data management plan (‘DMP’) is established and regularly updated. The use of this template is recommended for Horizon Europe beneficiaries. In completing the sections of the template the requirements for research data management of Horizon Europe as described in article 17 and analysed in the Annotated Grant Agreement, article 17, must be addressed.

	Title	UCPH Guidance
1.	**1. Data Summary**	
1.1	Will you re-use any existing data and what will you re-use it for? State the reasons if re-use of any existing data has been considered but discarded.	Indicate to which extent you will reuse research data that have been generated or collected in a previous project, by external partners, or that are available in a public or commercial database, repository, archive, from literature, homepages, social media, etc. State, whether you have conducted a systematic search of existing research data that might be relevant for reuse in the project (or if you are planning to so during the project). The following tools and resources may help you find available data: • Google Dataset Search: https://datasetsearch.research.google.com/ • DataCite Commons: https://commons.datacite.org/ • Open Data Sources: https://kub.kb.dk/datalab/opendata Note that existing data may underlie copyright restrictions or specific terms and conditions for their access and reuse.
1.2	What types and formats of data will the project generate or re-use?	Give an overview of the research data that you will collect or generate and use in the project. Research data can be quantitative and/or qualitative. Here, you should at least describe any research data that are in digital form. Additionally, you can include physical objects, software, models. etc. (see also Section 3). You should not include 'traditional' publications in journals or books and administrative project documentation. Group the research data according to different types or as distinct data sets, e.g. by their time, place and method of collection, their overall characteristics, or by work package. If known, specify the file formats. Provide a meaningful name / title or identifier for each data type / data set that you use as reference throughout the remaining document. Identify and indicate any data that contain personal, confidential or otherwise sensitive information and that require special protection.
1.3	What is the purpose of the data generation or re-use and its relation to the objectives of the project?	Describe how the different data types / data sets defined in the previous question will be used throughout the project (e.g. as input for a specific analysis or to generate the results for a particular publication). Clarify dependencies between different data types / data sets (e.g. if one data type / data set will be derived from another one or if two different data types / data sets will be combined). Where relevant, refer to the objectives and deliverables in the project plan and relate to individual work packages.

1.4	What is the expected size of the data that you intend to generate or re-use?	For each data type / data set, give an estimate on their expected amount and volume (e.g. number of files and file sizes). If that is unknown, indicate the scope of the data collection (e.g. number of repeated measurements or simulations, number of participants in a field study, time-span and frequency of observations, etc.). Note that the management of big data might require special infrastructure and generate additional costs.
1.5	What is the origin/provenance of the data, either generated or re-used?	Refer to relevant documentation, methods, software and equipment used to collect or generate the different data types / data sets. For existing data, state where and how the data can be accessed and retrieved and which persons or organizations are responsible for them. If relevant, describe how you can assess the validity, integrity and completeness of those data.
1.6	To whom might your data be useful ('data utility'), outside your project?	The research data collected or generated in your project can have significant value beyond the scope of the project and should be regarded as original output worth disseminating. Indicate potential target groups for the different data types / data sets, e.g. those that might be able to reuse the data for new research, innovation, decision-making, education or public awareness raising. Those can include other researchers in the same or other disciplines, interest groups, policy makers, commercial entities, media, etc. Keep your target groups and possible applications for reuse of the research data in mind when addressing what you will share and how in the next section.
2.	**2. FAIR data**	
2.1	**2.1. Making data findable, including provisions for metadata:** Will data be identified by a persistent identifier?	A persistent identifier provides a stable and permanent link to a resource that is available online. Examples for persistent identifiers are: • Digital Object Identifier (DOI) • Handle • Archival Resource Key (ARK) • Unique record identifier (UID) • Database accession number For each data type / data set, indicate whether you will deposit the data in an online repository, database or archive that issues persistent identifiers for stored items (see also Question 5 in this Section).

2.2	**2.1. Making data findable, including provisions for metadata:** Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.	Metadata provide information about the content, structure and context of research data. For each data type / data set, describe the metadata that you will create. General metadata for research data include: • Names of creators, authors and contributors • Collection dates and locations • Descriptive keywords • Definitions for abbreviations, parameters, units, etc. Metadata standards provide unambiguous and universal ways of describing research data and help others to interpret the data in the originally intended manner. Whenever possible, adhere to metadata standards that are broadly accepted and used in your research discipline. The Metadata Standards Catalog from the Research Data Alliance lists common examples: https://rdamsc.bath.ac.uk/
2.3	**2.1. Making data findable, including provisions for metadata:** Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?	For each data type / data set, specify meaningful keywords that describe the area of research and possible applications for reuse. Indicate, whether you will publish those keywords as part of the metadata in an online repository, database, archive or catalogue (see also next question). In order to increase the visibility and discoverability of your research data, you may also consider adding a short description of your research data and metadata to your researcher profile, the project's homepage or other relevant channels for outreach.
2.4	**2.1. Making data findable, including provisions for metadata:** Will metadata be offered in such a way that it can be harvested and indexed?	For each data type / data set, indicate whether you will publish associated metadata in an online repository, database, archive or catalogue that allows for searching and browsing the metadata and that exposes the metadata to common internet search engines. Check the specifications of the chosen repository, database, archive or catalogue on whether and how metadata are indexed and harvested.
2.5	**2.2. Making data accessible - Repository:** Will the data be deposited in a trusted repository?	Research data repositories are considered trustworthy when they as minimum: • Can ensure the online availability of deposited data and metadata for at least five years. • Are recognized and used by other researchers (in the same field). • Provide clear terms and conditions for providers and users of deposited data and metadata. • Are run by an active, sustainable and reputable organization. Trusted repositories also have to support the FAIR principles and must therefore: • Issue persistent identifiers (see Questions 1 and 7 in this Section). • Adhere to common metadata standards (see Question 2 in this Section). • Make the metadata findable online (see Question 4 in this Section). • Make the data accessible online (see Question 10 in this Section). • Provide a standard reuse license (see Question 21 in this Section).

		For each data type / data set, indicate, whether you will upload the data to a trusted repository and refer to a general description of that repository (e.g. the repository's homepage or a related publication). The Registry of Research Data Repositories (re3data) provides a very extensive and comprehensive overview of both general and discipline-specific repositories and their main features: http://www.re3data.org/
2.6	**2.2. Making data accessible - Repository:** Have you explored appropriate arrangements with the identified repository where your data will be deposited?	Indicate any special terms and conditions that apply for the deposition of data and metadata in any of the repositories identified in the previous question. Those can include: • Membership requirements • Upload or usage fees • Rights to and responsibilities for deposited data and metadata If relevant, describe how the selected repositories facilitate specific needs regarding e.g. security, access management and long-term preservation.
2.7	**2.2. Making data accessible - Repository:** Does the repository ensure that the data is assigned an identifier? Will the repository resolve the identifier to a digital object?	See also Question 1 in this Section about persistent identifiers. Research data repositories typically issue identifiers automatically upon deposition of data and metadata. Most identifiers serve as links that resolve directly to a unique landing page for the data and metadata in the repository. Refer to the specifications of the chosen repositories.
2.8	**2.2. Making data accessible - Data:** Will all data be made openly available? If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why, clearly separating legal and contractual reasons from intentional restrictions. Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if opening their data goes against their legitimate interests or other constraints as per the Grant Agreement.	For each data type / data set, indicate whether you will make the data openly available to others – that means free of charge, without access restrictions and with as little legal and technical barriers as possible. Note that Horizon Europe grant holders are expected to make their data openly available by default and must state any reasons for keeping data closed or for restricting access to data. Legitimate reasons include: • Legal, ethical or contractual obligations • Matters of security or confidentiality • Commercial interests • Copyright restrictions for existing data Personal data may only be made available according to the rules set out in the General Data Protection Regulation and the Danish Data Protection Act: https://kunet.ku.dk/work-areas/research/data/personal-data/disclosure/ Contact the Tech Transfer Office for help with the dissemination of research data underlying cooperation agreements, and for questions regarding commercialization of results, patents and Intellectual Property Rights: https://kunet.ku.dk/work-areas/research/innovation_and_business/

2.9	**2.2. Making data accessible - Data:** If an embargo is applied to give time to publish or seek protection of the intellectual property (e.g. patents), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.	For each data type / data set, indicate when you will make the data available to others. Note that you may publish a description of the data and corresponding metadata even before releasing the actual data. Some research data repositories allow for applying embargo periods, during which public access to deposited files is restricted or blocked.
2.10	**2.2. Making data accessible - Data:** Will the data be accessible through a free and standardized access protocol?	For each data type / data set that you will make available (either openly or restricted), describe how others will be able to retrieve the data. Data deposited in a repository will typically be available to others for direct download (from a web portal and/or through API's).
2.11	**2.2. Making data accessible - Data:** If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?	For each data type / data set that you will not make openly available, describe how you will enable access to authorized users, e.g. by defining roles and permissions for folders and files on a shared network drive. Some research data repositories allow for restricted access to deposited files, meaning that access can be granted to individual users or by request only. Note that any restrictions will require a permanent contact person managing access requests also after the end of the project.
2.12	**2.2. Making data accessible - Data:** How will the identity of the person accessing the data be ascertained?	For each data type / data set that you will not make openly available, define who will be authorized to access the data and under which conditions. If relevant, describe procedures and measures for the authorization and authentication of those users (e.g. additional documentation, registration and logging activities).
2.13	**2.2. Making data accessible - Data:** Is there a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data)?	For each data type / data set that you will not make openly available, define who will manage access requests both during and after the project. Note that special rules and conditions may apply when sharing personal, confidential or otherwise sensitive data, and that specific permissions or agreements may be required. Please review the guidelines on disclosure and sharing of personal data on the Research Portal: https://kunet.ku.dk/work-areas/research/data/personal-data/disclosure/Pages/default.aspx

2.14	**2.2. Making data accessible - Metadata:** Will metadata be made openly available and licenced under a public domain dedication CC0, as per the Grant Agreement? If not, please clarify why. Will metadata contain information to enable the user to access the data?	For each data type / data set that you will make available (either openly or restricted) in an online repository, database, archive or catalogue, indicate whether the associated metadata will be openly available (see also Question 4 in this Section about 'publication of metadata'). When depositing data in a research data repository, the associated metadata are typically made openly available by default (under a public domain dedication, e.g. CC0). Refer to the specifications of the chosen repositories for details. For restricted data, the description of the data and associated metadata should include the terms and conditions and instructions for requesting access.
2.15	**2.2. Making data accessible - Metadata:** How long will the data remain available and findable? Will metadata be guaranteed to remain available after data is no longer available?	For each data type / data set, specify the expected retention period of the data (after the end of the project). If certain data have to be discarded or deleted, indicate whether a description of those data and associated metadata will remain available. Where applicable, refer to the specifications of the chosen online repository, database, archive or catalogue on how long deposited data and metadata remain available. Note that according to UCPH's Policy for Research Data Management, data underlying publications should be available for a minimum of 5 years, see: https://research.ku.dk/integrity/documents/UCPH_Policy_for_Research_Data_Management_2022.pdf
2.16	**2.2. Making data accessible - Metadata:** Will documentation or reference about any software be needed to access or read the data be included? Will it be possible to include the relevant software (e.g. in open source code)?	For each data type / data set that uses non-standard or non-open formats, describe relevant documentation and indicate specific software (and version) and/or equipment required to open and process the data.
2.17	**2.3. Making data interoperable:** What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and re-use within and across disciplines? Will you follow community-endorsed interoperability best practices? Which ones?	For each data type / data set, refer to e.g. standards for metadata (including vocabularies, taxonomies and ontologies), file naming conventions, common formats, protocols, templates, standard procedures and best practices for documentation and dissemination of research data in your field. For metadata standards, refer also to Question 2 in this Section. • Vocabularies provide unambiguous definitions for names, variables, parameters and other metadata. • Taxonomies structure names, variables, parameters and other metadata in a hierarchy. • Ontologies include definitions for the possible relations between names, variables, parameters and other metadata.

2.18	**2.3. Making data interoperable:** 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? Will you openly publish the generated ontologies or vocabularies to allow reusing, refining or extending them?	Rigorously defined metadata standards such as controlled vocabularies and ontologies are prerequisite for machine-interoperability and most data science applications. If applicable, describe measures to make data and metadata machine-interoperable, e.g. for automatic harvesting, analysis, machine-learning and combination with other types of data.
2.19	**2.3. Making data interoperable:** Will your data include qualified references to other data (e.g. other data from your project, or datasets from previous research)? A qualified reference is a cross-reference that explains its intent. For example, X is regulator of Y is a much more qualified reference than X is associated with Y, or X see also Y. The goal therefore is to create as many meaningful links as possible between (meta)data resources to enrich the contextual knowledge about the data. (Source: https://www.go-fair.org/fair-principles/i3-metadata-include-qualified-references-metadata/).	Where applicable, metadata should include references to related types of data, publications, documentation, software, etc. Qualified references include information on the type of relation. Whenever possible, you should use persistent identifiers (e.g. DOI's) as links in the references.
2.20	**2.4. Increase data re-use:** How will you provide documentation needed to validate data analysis and facilitate data re-use (e.g. readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement, etc.)?	For each data type / data set that you will make available to others, specify relevant documentation on where and how the data have been collected, processed and analyzed. Describe how the documentation will be made available to others (e.g. within the data, as part of the description of the data, in the associated metadata, in a ReadMe file or as references to separate publications).

2.21	**2.4. Increase data re-use:** Will your data be made freely available in the public domain to permit the widest re-use possible? Will your data be licensed using standard reuse licenses, in line with the obligations set out in the Grant Agreement?	By default, (digital) research data collected or generated in Horizon Europe projects must be made available under a Creative Commons Public Domain Dedication (CC0) waiver, or Creative Commons Attribution (CC BY 4.0) usage license (or equivalent): • CC0 is a copyright waiver that allows others to "copy, modify, distribute and perform the work, even for commercial purposes, all without asking permission." • CC BY 4.0 allows others to freely share and adapt the work as long as they "give appropriate credit, provide a link to the license, and indicate if changes were made". For each data type / data set that you will make available to others, indicate the waiver or usage license that you will apply. For data that you will deposit in an online repository, database or archive, refer to the available options for applying waivers and usage licenses within the repository, database or archive. Justify any reasons for making data available under a more restrictive usage license or under non-standard terms and conditions.
2.22	**2.4. Increase data re-use:** Will the data produced in the project be useable by third parties, in particular after the end of the project?	For any data type / data set that you will make available to others but not in the public domain or with a standard usage license, describe the terms and conditions for reuse.
2.23	**2.4. Increase data re-use:** Will the provenance of the data be thoroughly documented using the appropriate standards?	Provenance includes relevant information on where, when, how and by whom data have been generated and edited. It helps others to understand the origin of the data and to assess their value for reuse. Where applicable, refer to standard methods, protocols, templates and procedures for collecting, processing and analyzing the data. Indicate, how you will keep track of changes to the data and versioning. Refer also to Question 20 in this Section on documentation.
2.24	**2.4. Increase data re-use:** Describe all relevant data quality assurance processes.	Quality assurance processes include measures and procedures for ensuring the integrity, completeness and validity of newly generated and existing data. Where applicable, refer to standards and best practices in your field.

3.	**3. Other research outputs**	
3.1	In addition to the management of data, beneficiaries should also consider and plan for the management of other research outputs that may be generated or re-used throughout their projects. Such outputs can be either digital (e.g. software, workflows, protocols, models, etc.) or physical (e.g. new materials, antibodies, reagents, samples, etc.).	While the main focus in Horizon Europe is on research data in digital form, you are encouraged to provide information about the management of other research output as well, such as physical materials, software and protocols (but not publications and administrative documents). Where applicable, you can add this information to the individual questions in the other sections. Otherwise, you can provide an overview here.
3.2	Beneficiaries should consider which of the questions pertaining to FAIR data above, can apply to the management of other research outputs, and should strive to provide sufficient detail on how their research outputs will be managed and shared, or made available for re-use, in line with the FAIR principles.	Outline, which other material and data (as described in the previous question) you will make available to others, and how. This may include: • Publication of open-source software, models and algorithms • (Pre-)registration of methods, reports and protocols • Registration of samples or artifacts in an online database • Data journal publications
4.	**4. Allocation of resources**	
4.1	What will the costs be for making data or other research outputs FAIR in your project (e.g. direct and indirect costs related to storage, archiving, re-use, security, etc.)?	Costs for managing research data can relate to infrastructure (e.g. for additional storage or increased security), licenses (e.g. for software or databases) and staff (data manager, programmer, trainer, etc.), depending on the types and amounts of material and data and the complexity of the project. These extensive guidelines from Utrecht University can help you estimate costs for research data management: https://www.uu.nl/en/research/research-data-management/guides/costs-of-data-management
4.2	How will these be covered? Note that costs related to research data/output management are eligible as part of the Horizon Europe grant (if compliant with the Grant Agreement conditions).	Refer to the Grant Agreement for eligible costs. Indicate how you will cover any costs for research data management in your project that are not covered or that exceed the funds provided through the grant.

4.3	Who will be responsible for data management in your project?	Different tasks for research data management may be assigned to different project members or designated support staff, including lab assistants, data managers, data stewards and curators, for example. Note that these might also have to be accounted for in the project plan and budget. In individual projects, the lead researcher will most likely be responsible for most of the activities outlined in the data management plan. However, you should clearly indicate tasks which require contributions by others. In larger projects and consortia, there will often be different roles assigned to different tasks or work packages, or at different locations, including technical and administrative support. Name the person(s) that will be responsible for e.g.: • Collection of data and material • Quality control and security • Analysis and documentation • Publication and curation of research data • Preservation Refer to project plans and other agreements, where relevant. For a general overview of possible roles in a research project, see the definitions for contributors in the CRediT taxonomy: https://credit.niso.org/
4.4	How will long term preservation be ensured? Discuss the necessary resources to accomplish this (costs and potential value, who decides and how, what data will be kept and for how long).	Long-term preservation of material and data might require additional provisions for securing their future accessibility and reusability. Your preservation strategy should thus address: • Long-term storage that sustains integrity and availability of the material and data as well as appropriate levels of security. • Durable formats that remain readable and/or executable. Where applicable, consider conversion of file formats and digitization of physical material. • Retention periods that reflect the significance of the material and data for future research, any external requirements and disciplinary traditions. • Associated documentation and metadata that enables others to fully understand the origin of the material and data and how they can be used. • Estimation of costs, including archival fees and efforts for future migration of data and conservation of materials, conversion of files, update of documentation and metadata, etc. • Assignment of responsibilities for published and archived research data after project end. These entail handling access requests and keeping documentation and contact information up to date. For recommendations regarding long term preservation and available tools at UCPH, see the Research Portal: https://kunet.ku.dk/work-areas/research/data/data-preservation/.

		Note also the requirements for long term preservation described in the UCPH Policy for Research Data Management: https://research.ku.dk/integrity/documents/UCPH_Policy_for_Research_Data_Management_2022.pdf
5.	**5. Data security**	
5.1	What provisions are or will be in place for data security (including data recovery as well as secure storage/archiving and transfer of sensitive data)?	Note that you might use different types of infrastructure solutions at different stages in the project and for different purposes (e.g. for data analysis, sharing and preservation) and that you should adapt the security provisions accordingly. Find an overview of UCPH's storage and sharing solutions on the Research Portal: https://kunet.ku.dk/work-areas/research/data/facilities-for-data-storage-and-sharing-in-active-projects/ Research projects containing personal data and biobanks must be registered in UCPH's joint record of biobanks and record of research projects containing personal data, using a specific registration form: https://kunet.ku.dk/work-areas/research/data/personal-data/personal-data-in-research-projects-and-biobanks/ Projects involving personal data must also carry out a risk assessment. In some cases with a high risk for data subjects, a Data Protection Impact Assessment (DPIA) is required as well. Find more information about risk and impact assessments on the Research Portal: https://kunet.ku.dk/work-areas/research/data/personal-data/impact-assessment/
5.2	Will the data be safely stored in trusted repositories for long term preservation and curation?	For each data type / data set that you will deposit in a research data repository (see Question 5 in Section 2), indicate how long the data will remain available. Specify all data types / data sets that you will preserve and curate for long term either parallel to or instead of deposition in a research data repository.
6.	**6. Ethics**	
6.1	Are there, or could there be, any ethics 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).	Legal, ethical and contractual obligations as well as any other concerns due to sensitivity, confidentiality and security might restrict the sharing of research data outside the project. Refer to ethical approvals and cooperation agreements as well as the project's ethics deliverables, where relevant. Find UCPH's guidelines and requirements for research involving personal data on the Research Portal: https://kunet.ku.dk/work-areas/research/data/personal-data/
6.2	Will informed consent for data sharing and long term preservation be included in questionnaires dealing with personal data?	Find information about informed consent to the processing of personal data under GDPR on the Research Portal: https://kunet.ku.dk/work-areas/research/data/personal-data/test-subjects/

7.	**7. Other issues**	
7.1	Do you, or will you, make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones (please list and briefly describe them)?	The UCPH Policy for Research Data Management applies to anyone conducting or supporting research activities at the University of Copenhagen: https://research.ku.dk/integrity/documents/UCPH_Policy_for_Research_Data_Management_2022.pdf Research projects containing personal data and research biobanks as well as biobanks must comply with the rules of the General Data Protection Regulation (GDPR): https://kunet.ku.dk/work-areas/research/data/personal-data/ Publishers may have policies on the availability and peer review of data underlying submitted manuscripts in place. Certain research disciplines (e.g. within bioinformatics, astronomy and social sciences) have established standards or best practices for the documentation and dissemination of data.