This template is intended for creating a data management plan, based on the data management section that was part of your research proposal. You are kindly requested to complete the plan below and submit it to CTFC. It is advised to regularly review the data management plan when required during the course of the research project.

|  |  |  |
| --- | --- | --- |
| **0** | **General Information** | |
| 0.1 | Dataset title and WP number | Provide the name of the dataset and the corresponding work package number. |
| 0.2 | Author and ORCID | Provide the full name of the creator/s of the dataset and the corresponding ORCID |
| 0.3 | Number of files | Indicate the number of data files included on the dataset. |
| 0.4 | Description | Describe the contents of the dataset, as well as the purpose, nature and scope. |
| 0.5 | Subject | Provide the dataset knowledge area  Agricultural Sciences  Arts and Humanities  Astronomy and Astrophysics  Business and Management  Chemistry  Computer and Information Science  Earth and Environmental Sciences  Engineering  Law  Mathematical Sciences  Medicine, Health, and Life Sciences  Physics  Social Sciences  Other |
| 0.6 | Keywords | Indicate the keywords that will be indexed for the purpose of retrieving the dataset (at least 5) |
| 0.7 | Data category | Categorize the data  Observational  Experimental  Simulated  Compiled |
| 0.8 | Type of data | Describe the data you expect your research will generate  Administrative records data  Aggregate data  Census/enumeration data  Clinical data  Coded documents  Coded textual  Compiled data  Encoded data  Event/transaction data  Experimental Data  Genomic data  Geospatial data  Laboratory notebook  Machine-readable text  Measurement and test data  Observation data/ratings  Process-produced  Program source code  Psychological test  Recorded data  Simulation data  Survey data  Textual data  Time budget diaries  Other (please specify) |
| 0.9 | Dataset stability | Categorize the stability of each dataset  Fixed dataset (never change after being collected or generated)  Growing dataset (new data may be added, but the old data is never changed or deleted)  Revisable dataset (new data may be added, and old data may be changed or deleted) |
| 0.10 | Related publication and DOI | Yes     No  Indicate if there is any related publication. If applicable, give the title and DOI of the related publication. |
| **1** | **What data will be collected or produced, and what existing data will be re-used?** | |
| 1.1 | Will you re-use existing data for this research? | Yes     No |
|  | **If yes**: explain which existing data you will re-use and under which terms of use. | Explain which existing data you will re-use and state any constraints on re-use of existing data if there are any. |
| 1.2 | Which file formats will be used for this dataset? | Give details on the data format: the way in which the data is encoded for storage.  Give preference to open and standard formats as they facilitate sharing and long-term re-use of data.  AutoCAD DXF version R12 (.dxf)  BWF (.bwf)  COLLADA (.dae)  CSV (.csv)  DICOM (.dcm)  FLAC (.flac)  GeoTIFF (.tif, .tiff)  GML (.gml)  HTML (.html)  JPEG (.jpg, .jpeg)  JPEG 2000 (.jp2)  JSON-LD  MATLAB (.)  Matroska (.mka)  MIF/MID (.mif, .mid)  MXF (.mxf)  NetCDF  NTriples (.nt)  ODS (.ods)  ODT (.odt)  OPUS  PDF/A (.pdf)  PNG (.png)  Polygon file format (.ply)  R  Raster GIS  RDF/XML (.rdf)  REFI-QDA (Qualitative Data Analysis)  SIARD(.siard)  SPSS (.dat/.sps)  SQL (.sql)  STATA (.dat/.DO)  SVG (.svg)  SVG (.svg)  TextFabric  TIFF (.tif, .tiff)  Trig (.trig)  Turtle (.ttl)  Unicode text (txt)  WaveFront Object (.obj)  X3D (.x3d)  XML (.xml)  Other (please specify) |
| 1.2 | How much data storage will your project require in total? | 0 – 10 GB      10 – 100 GB  100 – 1000 GB  >1000 GB |
| **2** | | **Who will have access to your datasets? How and when will you share your datasets, if applicable?** | |
| 2.1 | | How will access be managed and by whom? | Describe who will have access to each dataset at each stage of the research. Indicate how will access be managed and by whom. |
| 2.2 | | Will you encrypt your datasets? | Yes     No |
| 2.3 | | How will your datasets be shared? | Describe how you will share each dataset, if applicable |
| 2.4 | | What software and hardware are needed to reuse your datasets? | Indicate the software and hardware needed to reuse our dataset. |
| 2.5 | | Do you plan for hardware and software obsolescence? | Yes     No |
| 2.7 | | What restrictions will you have on data sharing (confidentiality, privacy, security, etc.)? | IMPORTANT: If you cannot share your dataset, reasons should be provided to justify your decision. Due consideration is given to aspects such as privacy, public security, ethical limitations, property rights and commercial interests. |
| **3** | **What metadata and documentation will accompany the data?** | |
| 3.1 | Indicate what documentation will accompany the data. | Consider what other documentation is needed to enable re-use. This may include information on the methodology used to collect the data, analytical and procedural information, definitions of variables, units of measurement, and so on.  Consider how this information will be captured and where it will be recorded, for example in a database with links to each item, a ‘readme’ text file, file headers, code books, or lab notebooks. |
| 3.2 | Indicate which metadata will be provided to help others identify and discover the data. | To be findable, accessible, interoperable, and reusable, data must be accompanied with descriptive information in the form of metadata.  Depositing data in a certified or trustworthy repository will typically involve providing information about the data according to a metadata standard scheme (typically Dublin Core). If this is the case for the data described in this plan, that can be specified here.  Use existing standards of the discipline when possible |
| **4** | **How will data and metadata be stored and backed up during the research?** | |
| 4.1 | Describe where the data and metadata will be stored and backed up during the project. | Institution networked research storage  Other (please specify)  Give preference to the use of robust, managed storage with automatic backup, such as provided by IT support services of your home institution.  Please specify if you make use of other storage solutions for storage and backup of research data during the project, in addition to or instead of the institutional research drive. This may be because you need more space than offered by your institution; to facilitate data sharing with collaborators; or because your data requires additional security. Please explain. |
| 4.2 | How will data security and protection of sensitive data be taken care of during the research?  Explanation: | Not applicable (no sensitive data)  Default security measures of the institution networked research storage.  Additional security measures (please specify)  Consider data protection, particularly if your data is sensitive – for example, containing personal data, politically sensitive information or information relating to religion and health, trade secrets or national security information. Describe the main risks and how these will be managed. |
| **5** | **How will you handle issues regarding the processing of personal information and intellectual property rights and ownership?** | |
| 5.1 | Will you process and/or store personal data during  your project? | Yes     No |
|  | **If yes**, how will compliance with legislation and (institutional) regulation on personal data be ensured? | If yes, it is strongly recommended that you seek advice from specialised support staff. You must ensure that when dealing with personal data, data protection laws. |
| 5.3 | How will ownership of the data and intellectual property rights to the data be managed? | Explain who will be the owner of the data, meaning who will have the rights to control access.  Make sure to cover these matters of rights to control access to data for multi-partner projects and multiple data owners in the consortium agreement.  Indicate whether intellectual property rights are affected. If so, explain which and how they will be dealt with. |
| 5.4. | How will you deal with privacy or confidentiality, if applicable? | Describe whether each dataset contains direct or indirect identifiers.  Describe whether consent to share the data will be gathered during the informed consent process.  Describe how shared data will be anonymized, if applicable. |
| **6** | **How and when will data be shared and preserved for the long term?** | |
| 6.1 | How will data be selected for long-term preservation? | All data resulting from the project will be preserved for at least 10 years  Other (please specify)  Indicate what data must be retained or destroyed for contractual, legal, or regulatory purposes. Indicate how it will be decided what data to keep. Describe the data to be preserved long-term.  It is expected to preserve the data resulting from your project for at least ten years, unless legal provisions or discipline-specific guidelines dictate otherwise. |
| 6.2 | Are there any (legal, IP, privacy related, security related) reasons to restrict access to the data once made publicly available, to limit which data will be made publicly available, or to not make part of the  data publicly available?  **If yes**, please explain. | Yes     No  Indicate whether there are any restrictions on the re-use of the data. If it is necessary to restrict access to certain parts of the data or to apply a data sharing agreement, explain how and why. Explain what actions will be taken to overcome or to minimize restrictions. |
| 6.3 | What data will be made available for re-use? | All data resulting from the project will be made available  Other (please specify)  Indicate what data will be made available for re-use. This selection may differ from the data that is preserved, when the data are so large that it is unfeasible to deposit the data in a repository in its entirety, or if there are reasons that prohibit making data available for re-use as specified in the previous question. |
| 6.4 | When will the data be available for re-use, and for how long will the data be available? | Explain when the data will be made available. Indicate the expected timely release. Explain the reason and duration of any embargo periods. Explain whether exclusive use of the data will be claimed and if so, why and for how long. Indicate whether data sharing will be postponed or restricted for example to publish, protect intellectual property, or seek patents.  Data available as soon as article is published  Data available upon completion of the project  Data available after completion of project (with embargo) |
| 6.5 | In which repository will the data be archived and made available for re-use? | CORA.RDR - Federated multidisciplinary repository for the publication of research datasets in FAIR mode following the European Open Science Cloud (EOSC) guidelines.  Additionally, other repositories (please specify) |
| 6.6 | How will you license your datasets? | Describe how each dataset will be licensed.  Data is not copyrightable. However, a presentation of data (such as a chart or table) may be.  Data can be licensed (please specify)  (CC0)  (CC BY-SA)  (CC BY-ND)  (CC BY-NC)  (CC BY-NC-SA)  (CC BY-NC-ND) |
| 6.7 | Describe your strategy for publishing the analysis software that will be generated in this project. | Indicate whether potential users need specific tools or software (specific scripts, codes or algorithms developed during the project, etc.) to access, interpret and (re-)use the data. Indicate how these items will be made available. Consider the sustainability of software needed for accessing and interpreting the data. |
| **7** | **Data management costs** | |
| 7.1 | What resources (for example financial and time) will be dedicated to data management and ensuring that data will be FAIR (Findable, Accessible, Interoperable, Re-usable)? | Explain how the necessary resources (for example time) to prepare the data for sharing/preservation (data curation) have been costed in. Indicate whether additional resources will be needed to prepare data for deposit or to cover any charges from data repositories. If yes, explain how much is needed and how such costs will be covered**.** |