# STFC DMP Template *(with guidance notes)*

### Admin Details

**Plan Name:** STFC Data Management Plan

**Principal Investigator / Researcher:**

**Funder:** STFC

**Institution:** Royal College of Art

### 1. Data types

**Specify the types of data the research will generate.**

***STFC Guidance*** *Data management plans should describe the types of data that are expected to be produced from the project, including the raw data arising directly from the research, the reduced data derived from it, and published data.*

***RCA/DCC guidance on Data Type*** *Questions to consider:*

* *What types of data will you create?*
* *Which types of data will have long-term value?*

*Outline the types of data that are expected to be produced from the project e.g. quantitative, qualitative, survey data, experimental measurements, models, images, audiovisual data, samples etc. Include the raw data arising directly from the research, the reduced data derived from it, and published data.*

### 2. Data preservation

**2.1 Specify which data will be preserved and how.**

***STFC Guidance*** *Unless there are compelling reasons not to do so, STFC expects data to be managed through an established repository, chosen to maximise the scientific value from aggregation of related data. This may be at the grant holder's institution or elsewhere. Data management plans may refer to the general policies of the chosen repository and only include further details if necessary to the specific project. (If it is proposed not to use an established repository, the data management plan will need to demonstrate that resources and systems will be in place to enable the data to be curated effectively beyond the lifetime of the grant, although STFC recognises that applicants may not have the expertise to describe in detail how data will be curated).*

***RCA/DCC guidance on Preservation Plan*** *Questions to consider:*

* *What is the long-term preservation plan for the dataset? e.g. deposit in a data repository*
* *Will additional resources be needed to prepare data for deposit or meet charges from data repositories?*

*Researchers should consider how datasets that have long-term value will be preserved and curated beyond the lifetime of the grant. Also outline the plans for preparing and documenting data for sharing and archiving.*

*If you do not propose to use an established repository, the data management plan should demonstrate that resources and systems will be in place to enable the data to be curated effectively beyond the lifetime of the grant.*

***RCA/DCC guidance: Data repository***

*Long-term preservation and access is generally best managed by using a specialist repository. While you don’t have to specify the repository you will use, you should state the criteria you will use to select it. When considering a repository, you should examine their policies, procedures, metadata standards and any costs that might be incurred. If using a storage facility other than an established repository or data centre, you will need to demonstrate its efficacy and longevity.*

*Some funders specify a data repository, such as*[*UK Data Service ReShare*](http://reshare.ukdataservice.ac.uk/)*,*[*NERC Data Centres*](http://www.nerc.ac.uk/research/sites/data/)*or*[*Archaeology Data Service*](http://archaeologydataservice.ac.uk/)*. Resources such as*[*re3data*](http://www.re3data.org/)*and those provided by* [*BBSRC*](https://bbsrc.ukri.org/research/resources/#datasharing)*or*[*Nature*](https://www.nature.com/sdata/policies/repositories) *can be used to find an appropriate repository. General purpose repositories that you may consider are* [*Zenodo*](https://zenodo.org/) *and* [*Figshare*](https://figshare.com/)*; these are non-discipline specific open access repositories that will ensure the preservation of data for a minimum of 10 years from the last point of access and provide a permanent DOI for the data. Alternatively, RCA researchers can deposit small datasets, particularly those containing textual or visual material, in the RCA Research Repository. All research data selected for long-term preservation should be registered in the RCA Research Data Repository, irrespective of where the data files themselves are deposited. Research data in non-digital formats, and digital data that cannot be made accessible or requires controlled access, should also be registered in the RCA Research Repository. This will increase the discoverability and visibility of the research data.*

**2.2 Specify the software and metadata implications.**

***STFC Guidance*** *The data management plan should specify the software and metadata that will be retained to enable the data to be read and interpreted.*

***RCA/DCC Guidance*** *Data documentation provides the information necessary to identify, understand and reuse your data. When this information is provided in a much more structured form it is known as 'metadata' (information about data). Without this information it may be impossible to understand or reuse the data.*

*Things to consider:*

1. ***What information about your data will you capture?***

*At a minimum your documentation should include project-level information such as details of who created or contributed to the data; how, why and when the data were created; description of the contents of the dataset; details of how and under what conditions the data can be accessed.*

*Where appropriate you should also include more data-specific information such as lists of variable names and definitions, values and their meanings, units of measurement, the representation of null values, descriptions of processing activities, software needed to access the data.*

1. ***What documentation will accompany your data?***

*Examples of data documentation include: research and laboratory notebooks, data dictionaries and codebooks, README txt files and descriptions of methods and protocols.*

*Consider also, whether there are other types of supporting documentation which could further help others to understand your data e.g. workshop or project diaries, blank consent forms, information sheet templates, survey tools, blank questionnaires/case report forms etc.*

*Consider using an existing metadata standard where such a standard exists. The Digital Curation Centre (DCC) maintains a list of*[*metadata standards*](http://www.dcc.ac.uk/resources/metadata-standards)*used in different disciplines.*

**2.3 Specify for how long the data will be preserved.**

***STFC Guidance*** *This may depend on the type of data. Where possible, STFC expects the original data, from which other related data can in principle be derived, to be retained for a minimum of 10 years from the end of the project. For data that by their nature cannot be re-measured, efforts should be made to retain them indefinitely.*

***RCA/DCC guidance:*** ***Preservation***

*Describe how you will preserve and share your data, including the length of time they will be kept and the nature of the storage location. The RCA Research Data Management Policy requires that all data needed to validate research findings are kept for a minimum of 10 years. Also indicate if any additional resources or funding will be required to deposit and store the data.*

*Funders generally expect data with long-term value to be preserved and remain accessible, alongside the software and code needed to reproduce your findings. This does not mean that you need to keep all of your data, but you will need to state who will be responsible for choosing and archiving data, as well as documenting the removal of any data that must be destroyed.*

*It is particularly important to preserve data which cannot be remeasured or recreated. Many research funders specify which data need to be preserved, how long for and where they should be deposited. See the DCC guide*[*How to appraise and select research data for curation*](http://www.dcc.ac.uk/resources/how-guides/appraise-select-data)*.*

### 3. Data sharing

**3.1 Specify and justify which data will have value to others and should be shared.**

***STFC Guidance*** *Any data that are shared should be of a sufficiently high quality to be of value to other researchers. In general, published data (data that are displayed or otherwise referred to in a publication) should be made publicly available, but it is for applicants to consider and justify which types of data will, in the context of their project, meaningfully and practically constitute published data. Publicly available means available to anyone, but there may be a requirement for registration to enable tracking of data use and to provide notification of terms and conditions of use where they apply. Other data should be made available wherever it is appropriate and cost-effective to do so, taking into account the cost of curation compared with the cost or feasibility of re-creation, the potential long-term demand for the data and the feasibility of their reuse by others.*

***RCA/DCC guidance:*** ***Data sharing***

*Outline which data you will share and how you will share them, e.g. depositing in a repository, using a secure data service or dealing with data requests individually. The method(s) used will depend upon the size and nature of the data. You should use standards and formats that enable reuse, and ensure data is discoverable through use of accurate metadata and persistent identifiers.*

*The Digital Curation Centre provides useful advice about*[*data appraisal and selection*](http://www.dcc.ac.uk/resources/how-guides/appraise-select-data)*.*

*Most funders allow a delayed release to allow researchers to have exclusive use of their data and to exploit the results of their research. See the RCA page on Research Funder Policies to determine when you need to make your data available. Restrictions on the release of data may be allowed, to protect confidentiality and for other ethical and legal reasons.*

*While restrictions on sharing should be minimised, you should take into account the following when sharing data:*

* *Does your data include confidential and sensitive information?*
* *Have participants given consent for their data to be shared?*
* *Consider what can be done to make sensitive data openly sharable - can these data be anonymised?*
* *Do different parts of your data require different access conditions? These may require separate deposits.*
* *Who will be responsible for controlling access?*

*Whatever form of publishing is used, research data should be licensed to indicate what users may or may not do with the data. Data repositories will indicate what licences are available for the data they house. More information is available from the Digital Curation Centre on* [*how to license research data*](http://www.dcc.ac.uk/resources/how-guides/license-research-data)*.*

*For all Royal College of Art research, a metadata record should be registered in the RCA Research Repository.*

*A Data Access statement should also be included in any publication based upon the research data. A Data Access Statement is a short statement explaining where the data is available, and under what license or access conditions. This helps to further increase the visibility of the data whilst also supporting the validity and reproducibility of your research findings.*

**3.2 Specify and justify the length of any proprietary period.**

***STFC Guidance*** *This might for example refer to the reasonable needs of the research team to have a first opportunity to exploit the results of their research, including any intellectual property arising. Where there are accepted norms within a scientific field or specific archive they should normally be followed. In general, STFC expects that published data should be made publicly available within six months of publication unless justified otherwise.*

***RCA/DCC guidance on Timeframe For Data Sharing*** *Questions to consider:*

* *When will you make the data available?*

*Data (with accompanying metadata) should be shared in a timely fashion. It is generally expected that timely release would be no later than publication of the main findings and should be in-line with established best practice in the field. Researchers have a legitimate interest in benefiting from their investment of time and effort in producing data, but not in prolonged exclusive use. Research funders typically allow embargoes in line with practice in the field, but expect these to be outlined up-front and justified.*

**3.3 Specify how data will be shared**

***STFC Guidance*** *The minimum level of data sharing expected would be that of making the data available in the natural format in which they were created, along with documentation and metadata, according to the standard accepted procedures within the scientific field. Where the data are likely to be in great demand by others it may be appropriate to request resources for a more proactive approach to data sharing, which maximises opportunities for cross linkage with other sectors.*

***RCA/DCC guidance on Method For Data Sharing***

*See the guidance on data sharing in section 3.1 above.*

### Resources

**Specify and justify any resources required to preserve and share the data.**

***STFC Guidance*** *Wherever possible, data management should make use of existing skills and capabilities. However, justification should be made for any additional specialist staff (or training for existing staff) needed within the grant to enable the research team to manage, preserve and share data effectively; and for any computational facilities needed to manage, store and share the data generated by the research.*

***RCA/DCC guidance on Resourcing*** *Questions to consider:*

* *Is additional specialist expertise (or training for existing staff) required?*
* *Do you require hardware or software which is additional or exceptional to existing institutional provision?*
* *Will charges be applied by data repositories?*

*Carefully consider any resources needed to deliver the plan. Where dedicated resources are needed, these should be outlined and justified. Outline any relevant technical expertise, support and training that is likely to be required and how it will be acquired. Provide details and justification for any hardware or software which will be purchased or additional storage and backup costs*

*.*

*Funding should be included to cover any charges applied by data repositories, for example to handle data of exceptional size or complexity. Also remember to cost in time and effort to prepare data for deposit and ensure it is adequately documented to enable reuse. If you are not depositing in a data repository, ensure you have appropriate resources and systems in place to share and preserve the data.*

*See UKDS guidance on* [*costing data management*](http://ukdataservice.ac.uk/manage-data/plan/costing.aspx)*.*