# MRC DMP Template *(with guidance notes)*

### Admin Details

**Plan Name:** MRC Data Management Plan

**Principal Investigator / Researcher:**

**Funder:** MRC

**Institution:** Royal College of Art

### 0. Proposal name

**0. Enter the proposal name**

***MRC guidance:***

*Exactly as in the proposal that the DMP accompanies*

### 1. Description of Data

**1.1 Type of Study**

***MRC Guidance***

*Up to three lines of text that summarise the type of study (or studies) for which the data are being collected.*

**1.2 Types of Data**

***MRC Guidance***

*Types of research data to be managed in the following terms: quantitative, qualitative; generated from surveys, clinical measurements, interviews, medical records, electronic health records, administrative records, genotypic data, images, tissue samples,...*

***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.*

**1.3 Format and scale of the data**

***MRC Guidance***

*File formats, software used, number of records, databases, sweeps, repetitions,... (in terms that are meaningful in your field of research). Do formats and software enable sharing and long-term validity of data?*

***RCA/DCC guidance on Data Volumes***

*Questions to consider:*

* *Do you have sufficient storage?*
* *Do you need to include costs for additional managed storage?*
* *Will the scale of the data pose challenges when sharing or transferring data between sites?*

*Consider the implications of data volumes in terms of storage, backup and access. Estimate the volume of data in MB/GB/TB and how this will grow to make sure any additional storage and technical support required can be provided.*

***RCA/DCC guidance on Data Format***

*Questions to consider:*

* *What format will your data be in?*
* *Why have you chosen to use particular formats?*
* *Do the chosen formats and software enable sharing and long-term validity of data?*

*Outline and justify your choice of format e.g. SPSS, Open Document Format, tab-delimited format, MS Excel. Decisions may be based on staff expertise, a preference for open formats, the standards accepted by data centres or widespread usage within a given community. Using standardised and interchangeable or open lossless data formats ensures the long-term usability of data.*

*See UKDS Guidance on* [*recommended formats*](http://ukdataservice.ac.uk/manage-data/format/recommended-formats.aspx)*.*

### 2. Data collection / generation

**2.1 Methodologies for data collection / generation**

***MRC Guidance***

*How the data will be collected/generated and which community data standards (if any) will be used at this stage.*

***RCA/DCC guidance on Data Capture Methods***

*Questions to consider:*

* *How will the data be created?*
* *What standards or methodologies will you use?*
* *How will you structure and name your folders and files?*
* *How will you ensure that different versions of a dataset are easily identifiable?*

*Outline how the data will be collected/generated and which community data standards (if any) will be used at this stage. Indicate how the data will be organised during the project, mentioning for example naming conventions, version control and folder structures. Consistent, well-ordered research data will be easier for the research team to find, understand and reuse.*

**2.2 Data quality and standards**

***MRC Guidance***

*How consistency and quality of data collection / generation will be controlled and documented, through processes of calibration, repeat samples or measurements, standardised data capture or recording, data entry validation, peer review of data or representation with controlled vocabularies.*

***RCA/DCC guidance on Data Quality***

*Questions to consider:*

* *How will you control data capture to ensure data quality?*
* *What quality assurance processes will you adopt?*

*Explain how the consistency and quality of data collection will be controlled and documented. This may include processes such as calibration, repeat samples or measurements, standardised data capture or recording, data entry validation, peer review of data or representation with controlled vocabularies.*

### 3. Data management, documentation and curation

**3.1 Managing, storing and curating data**

***MRC Guidance***

*Briefly, how data will be stored, backed-up, managed and curated in the short to medium term. Specify any community agreed or other formal data standards used (with URL references). [Enter data security standards in Section 4].*

***RCA/DCC guidance: Storage & security***

*Describe where the data will be stored and backed up during the course of research activities. This may vary if you are doing fieldwork or working across multiple sites so explain each procedure.*

*Identify who will be responsible for backup and how often this will be performed. The use of robust, managed storage with automatic backup, for example, that provided by university IT teams, is preferable. Storing data on laptops, computer hard drives or external storage devices alone is very risky.*

*See UK Data Service Guidance on*[*data storage*](https://www.ukdataservice.ac.uk/manage-data/store)*or DataONE Best Practices for*[*storage*](https://www.dataone.org/best-practices/storage)*.*

*Also consider data security, particularly if your data is sensitive e.g., detailed personal data, politically sensitive information or trade secrets. Note the main risks and how these will be managed. Also note whether any institutional data security policies are in place.*

*Identify any formal standards that you will comply with, e.g., ISO 27001. See the DCC Briefing Paper on Information Security Management -*[*ISO 27000*](http://www.dcc.ac.uk/resources/briefing-papers/standards-watch-papers/information-security-management-iso-27000-iso-27k-s)*and UK Data Service guidance on*[*data security*](https://www.ukdataservice.ac.uk/manage-data/store/security)*.*

*The RCA makes available the institutionally managed Google Drive suite of applications. Google Drive data is stored on servers within the EU and has been assessed by the RCA as a safe and appropriate venue for research data. Google Drive allows for files and data to be accessed from multiple device, so multiple project team members can work on them collaboratively. Google Drive also permits individual permissions so access to sensitive data can be managed as appropriate with internal and external partners. As a Cloud-based online technology, Google Drive removes the risk of data loss as automatic backup of all data is ensured. Furthermore, Google Suite has in-built version control meaning that older versions of the data are retained and backed up, thus guarding against human input error and ensuring retrieval of older versions if necessary. Google Suit undergoes regular independent audits on their data centres, network and operations. This is in compliance with the certified industry standards such as ISO 27001 and 27017.*

**3.2 Metadata standards and data documentation**

***MRC Guidance***

*Plans for documenting, annotating and describing data so that research data are usable by others than your own team. This may include documenting the methods used to generate the data, analytical and procedural information, capturing instrument metadata alongside data, documenting provenance of data and their coding, detailed descriptions for variables, records, etc.*

***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.*

**3.3 Data preservation strategy and standards**

***MRC Guidance***

*Plans and place for long-term storage, preservation and planned retention period for the research data. Formal preservation standards, if any. Indicate which data may not be retained (if any).*

***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)*.*

***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.*

### 4. Data security and confidentiality of potentially disclosive personal information

**4.1 Formal information/data security standards**

***MRC Guidance***

*Identify formal information standards with which your study is or will be compliant. An example is ISO 27001.*

***RCA/DCC guidance on Data Security***

*Questions to consider:*

* *What are the risks to data security and how will these be managed?*
* *Will you follow any formal standards?*

*If your data is sensitive (e.g. detailed personal data, politically sensitive information or trade secrets) you should discuss any appropriate security measures that you will be taking. Note the main risks and how these will be managed. Identify any formal standards that you will comply with e.g. ISO 27001.*

*See DCC Briefing Paper on* [*Information Security Management - ISO 27000*](http://www.dcc.ac.uk/resources/briefing-papers/standards-watch-papers/information-security-management-iso-27000-iso-27k-s)*.*

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

**4.2 Main risks to data security**

***MRC Guidance***

*If not using formal standards, summarise the main risks to the confidentiality and security of information related to human participants, and how these risks will be managed. Cover the main processes or facilities for storage and processing of personal data, data access, with controls put in place and any auditing of user compliance with consent and security conditions.*

*MRC guidance on the categories of data availability is provided.*

### 5. Data sharing and access

**5.1 Suitability for sharing**

***MRC Guidance***

*Indicate whether the data you propose to collect (or existing data you propose to use) in the study will be suitable for sharing ('Yes' or 'No')*

*If no, indicate why they will not be suitable for sharing and then go to Section 6.*

**5.2 Discovery by potential users of the research data**

***MRC Guidance***

*Indicate how potential new users can find out about your data and identify whether they could be suitable for their research purposes, e.g. through summary information (metadata) being readily available on the study website, in the MRC gateway for population and patient research data, or in other databases or catalogues. Indicate whether your policy or approach to data sharing is (or will be) published on your study website (or by other means).*

*Identify any data repository (-ies) that are, or will be, entrusted with storing, curating and/or sharing data from your study, where they exist for particular disciplinary domains or data types. Information on repositories is available here.*

***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.*

**5.3 Governance of access**

***MRC Guidance***

*Identify who makes or will make the decision on whether to supply research data to a potential new user.*

*For population health and patient-based research, indicate how* [*independent oversight of data access and sharing*](http://www.mrc.ac.uk/news-events/publications/mrc-policy-and-guidance-on-sharing-of-research-data-from-population-and-patient-studies) *(please see page 10 of PDF file generated by selecting the above or adjacent*[*link*](http://www.mrc.ac.uk/news-events/publications/mrc-policy-and-guidance-on-sharing-of-research-data-from-population-and-patient-studies/)*) works (or will work) in compliance with* [*MRC policy*](http://www.mrc.ac.uk/documents/pdf/mrc-data-sharing-policy/)*.*

*Indicate whether the research data will be deposited in and available from an identified community database, repository, archive or other infrastructure established to curate and share data.*

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

*Questions to consider:*

* *How will you make the data available to others?*
* *With whom will you share the data, and under what conditions?*

*Consider where, how, and to whom the data should be made available. Will you share data via a data repository, handle data requests directly or use another mechanism?*

*The methods used to share data will be dependent on a number of factors such as the type, size, complexity and sensitivity of data. Mention earlier examples to show a track record of effective data sharing.*

**5.4 The study team's exclusive use of the data**

***MRC Guidance***

*MRC's requirement is for timely data sharing, with the understanding that a limited, defined period of exclusive use of data for primary research is reasonable according to the nature and value of the data, and that this restriction on sharing should be based on simple, clear principles.*

*Summarise the principles of your current/intended policy.*

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

*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?*

**5.5 Restrictions or delays to sharing, with planned actions to limit such restrictions**

***MRC Guidance***

*Restriction to data sharing may be due to participant confidentiality, consent agreements or IPR. Strategies to limit restrictions may include data being anonymised or aggregated; gaining participant consent for data sharing; gaining copyright permissions. For prospective studies, consent procedures should include provision for data sharing to maximise the value of the data for wider research use, while providing adequate safeguards for participants. As part of the consent process, proposed procedures for data sharing should be set out clearly and current and potential future risks associated with this explained to research participants.*

***RCA/DCC guidance on Restrictions on Sharing***

*Questions to consider:*

* *Are any restrictions on data sharing required? e.g. limits on who can use the data, when and for what purpose.*
* *What restrictions are needed and why?*
* *What action will you take to overcome or minimise restrictions?*

*Outline any expected difficulties in data sharing, along with causes and possible measures to overcome these. Restrictions to data sharing may be due to participant confidentiality, consent agreements or IPR. Strategies to limit restrictions may include: anonymising or aggregating data; gaining participant consent for data sharing; gaining copyright permissions; and agreeing a limited embargo period.*

**5.6 Regulation of responsibilities of users**

***MRC Guidance***

*Indicate whether external users are (will be) bound by data sharing agreements, setting out their main responsibilities.*

***RCA/DCC guidance on Managed Access Procedures***

*Questions to consider:*

* *Will access be tightly controlled or restricted? e.g. by using data enclaves / secure data services*
* *Will a data sharing agreement be required?*
* *How will the data be licensed for reuse?*

*Indicate whether external users will be bound by data sharing agreements, licenses or end-user agreements. If so, set out the terms and key responsibilities to be followed. Note how access will be controlled, for example by the use of specialist services. A data enclave provides a controlled secure environment in which eligible researchers can perform analyses using restricted data resources. Where a managed access process is required, the procedure should be clearly described and transparent.*

### 6. Responsibilities

***MRC Guidance***

*Specify who, alongside the PI, is responsible for ensuring the study-wide data management, as well as for specific roles such as metadata creation, data security and quality assurance of data.*

***RCA/DCC Guidance*** *Questions to consider:*

* *Who is responsible for implementing the DMP, and ensuring it is reviewed and revised?*
* *Who will be responsible for each data management activity?*
* *How will responsibilities be split across partner sites in collaborative research projects?*
* *Will data ownership and responsibilities for RDM be part of any consortium agreement or contract agreed between partners?*

*Outline the roles and responsibilities for all activities e.g. data collection, metadata production, data quality, storage and backup, data curation & data sharing. Consider who will be responsible for ensuring relevant policies will be respected. Individuals should be named where possible.*

### 7. Relevant policies

**Relevant institutional, departmental or study policies on data sharing and data security**

### 8. Author and contact details

**Author of this Data Management Plan (Name) and, if different to that of the Principal Investigator, their telephone & email contact details**