Data Management Planning

AHRC funding applicants

Version 2.5 June 2021

University of Bristol

Research Data Service

Image: Marcus Aurelius Denarius2.jpg, Wikipedia, CC-BY-SA 3.0

SUMMARY

- From 29th March 2018, *all* AHRC applications for Leadership Fellows, Research Grants and Follow-on Funding must include a two-page Data Management Plan (DMP). This replaces the Technical Plan.

- Data Management Plans will be assessed by members of the Academic College and reviewed holistically as part of the whole proposal.

- Research data should be stored effectively, both during and after the project, and shared for a minimum of three years post-project.

- Any ethical and legal issues should be considered during the collection, storage, and release of data.

INTRODUCTION

In line with other UKRI funding councils, the AHRC is committed to the principle that those who receive research funding should take responsibility for the duration, management, and exploitation of their research data for future use. From the 29th March 2018, applicants are expected to explain how they will do this by completing a *Data Management Plan* at the time of application. Your Data Management Plan must cover the creation of, management of, and access to research data that results from your AHRC-funded research.

The Data Management Plan replaces the previous Technical Plan, which was assessed separately by a technical reviewer and only required where a project expected to produce significant digital outputs. The new Data Management Plan, along with any associated data management costs, is mandatory for *all* Leadership Fellows, Research Grants and Follow-on-Funding applications (but is not required for Research Networking) and should be added to the Je-S form as an attachment. Digital schemes which attract specific applications will have their own bespoke attachment types as determined by the requirements of the scheme. The AHRC have produced a list of FAQs to explain the reasons for the changes from a Technical Plan to a Data Management Plan.[1]

THE AHRC DATA MANAGEMENT PLAN (DMP)

Your DMP should explain how you'll manage any research data that you plan to use or create.[2] An assessment of the DMP will be made by members of the Academic College as part of the general assessment of your application (if a proposal is inherently technical in nature reviewers will be selected who are appropriate to review such proposals). Reviewers will be asked to consider whether the DMP seems feasible, sensible, appropriate and valid. A poorly prepared DMP may have a detrimental effect on an otherwise strong application.

Your DMP should be no longer than two pages and address the following points: -

---

[1] 'DMP FAQ'
https://ahrc.ukri.org/documents/guides/dmp-frequently-asked-questions/

[2] AHRC Data Management Plan
https://ahrc.ukri.org/documents/guides/data-management-plan/

1. Types of data the research will create – briefly introduce the data that will be produced and why you decided to use these data types.

2. Proposed methodologies – details of methodologies that will be used to create the data; information on why the project team selected will be suitable for the digital aspects of the work; details of how the institution's data support teams may need to support the project.

3. Data storage (short term) - how will the data be stored in the short term; backup plans during the project to ensure no data is lost.

4. Data storage (long term) – where data will be stored and why the choice is appropriate; how long it will be stored for and why; any costs associated with long-term storage.

5. Data sharing – the value data will have to others and how it could be used in the future; how the data will be shared; when data will be released and justification; how different audiences will be informed; any required updates to data; will data be open and any charges for access; any financial implications of sharing.

6. Ethical and Legal considerations – in collecting, storing and releasing the data.

If technical research issues such as these (often collectively known as Research Data Management) are entirely new to you, you may find it helpful to complete our 'Research Data Bootcamp' before going any further.[3]

---

[3] Research Data Bootcamp,
https://data.bris.ac.uk/bootcamp/

Types of data

Use this section to briefly introduce the types of data the project will create. If there are several different types, providing this information as a list or table can be helpful. Research data is essentially information that you use as an integral part of your research. Digital research data can be regarded as that created in a digital form (born digital) or converted to a digital form (digitized). Examples can include interview recordings, transcripts, videos, scans, photographs and databases. Research data does not include incidental or administrative data generated in the course of personal activities, desktop or mailbox backups, or data produced by non-research activities such as University administration or teaching.

As part of your DMP you should state in which format(s) your data will be collected, analysed and stored (for example, Open Document Format, .CSV file or Excel spreadsheet). A significant barrier to sharing any research digitally, and one you should address in your DMP, is the widespread use of highly specialised file formats. In order to use any digital file, a number of digital technologies must be available, which are known as technological 'dependencies'. These may be fairly common technologies such as a desktop PC, the Windows 10 operating system, and Adobe Reader DC 15 software. Or, the technology required might be rare and hard to acquire, or even unique (for example Windows 95 or any software package made by a single vendor).

You should address this problem by minimising the number of technological dependencies involved in using your digital output/technology as much as possible. Where dependencies are inevitable you should favour 'open' technologies rather than proprietary ones. Proprietary technologies are owned by a vendor or group of vendors. Commercial pressures may lead to the withdrawal of a particular piece of commercial hardware or software, in favour of a new and possibly incompatible replacement. In contrast, 'open' technologies are supported by a community of users and do not have the same commercial vulnerabilities.

When selecting a file format, your own research needs must come first. If you find you need to use an unusual or non-standard format (one that isn't widely used) you should consider converting it into a more widely re-usable format once you have finished using the data yourself. If you're unsure which file formats are 'open' and/or widely used, the UK Data Archive publishes a list of recommended deposit formats[4] that they are prepared to accept. This list should give you an insight into which file formats are likely to have longevity.

It is also useful in your DMP to try and estimate the size of the data you expect to generate. This can be difficult to do before a study begins; if necessary, use quantities generated in a similar previous study as the basis for your estimate.

Data creation and proposed methodologies

This section is all about describing how the data you listed in the previous section will be gathered or created, and the methodologies you plan to use to do this. So, for example, if you are conducting interviews you might be going into the field and using a voice recorder to make sound files of the interview, and then later producing transcriptions. Also make sure the plans that you describe correspond with the tasks outlined in the Case for Support.

In this section, you should provide information about the technical expertise of anyone who will be part of your proposed project. Identify the responsibilities of key individuals if possible (for example: "The Research Assistant will be responsible for ensuring the video recordings are copied to RDSF storage as soon as possible after they have been made."). Also, mention any institutional data support teams that you intend to call on. This may include Research IT[5] (research software development), ACRC[6] (research data storage) or the Communications and Marketing Office[7] (project websites).

Data storage

The next two sections of the DMP deal with data storage, both whilst the project is actually taking place, and also the arrangements for the long-term storage of the data beyond the life of the project itself.

[4] UK Data Archive File Formats Table, https://www.ukdataservice.ac.uk/manage-data/format/recommended-formats
[5] Research IT https://www.bristol.ac.uk/research-it/

[6] Advanced Computing Research Centre, https://www.acrc.bris.ac.uk/
[7] Public Relations https://www.bristol.ac.uk/communications-marketing/

Short-term storage

This part of the DMP should focus on the backup arrangements you will have in place whilst the project is running, to ensure no data is lost. It is recommended that, as data is created, University of Bristol staff store it in the Research Data Storage Facility (RDSF), managed by the Advanced Computing Research Centre (ACRC).[8] Each member of research staff is entitled to 5TB of storage, without charge. Contact ACRC directly to request your allocation.

If you have used up your storage quota, or your project requires more storage space than the initial allocation, there will be a cost, and ACRC should be contacted before your application is finalised. The back-up procedures, policies and controlled access arrangements used by the RDSF are of a very high standard and RDSF can provide a description of them to go into your application. If you do not intend to make use of RDSF, your data storage provider's back-up procedures should be described instead.

You should explain how you'll keep your data safe before it's deposited in a secure storage facility (such as the RDSF). This is particularly important if you're conducting field research. As a minimum, try to ensure that at all times more than one copy of the data exists and that every copy can easily be accounted for and located if required.

Long-term storage

This section is concerned with the long-term preservation of your research data beyond the life of

the project. You should describe in this section where you plan to store your data once the project has finished, why you have chosen this storage option, and for how long data will be stored. By far the best option is to use the RDSF for this purpose. Data is securely stored for a minimum of 20 years, which more than meets the three-year post-project requirement of AHRC.

It is important to describe any costs associated with the long-term storage of your data and include these in the Justification for Resources. AHRC awards cannot cover data management costs incurred after the end of a grant, but they can be used for costs incurred *within* the lifetime of a project. If your storage requirements exceed 5TB there will be a charge for using the RDSF, but this is an upfront "pay once, store forever" cost which can be included in your application.

If you do not intend to make use of RDSF, your storage provider's back-up procedures should be described instead. If you will be working collaboratively with other institutions, make sure that the security and back-up procedures of each data holding partner are outlined within the DMP.

Data sharing

This is a significant part of the DMP, as it is here that you are expected to describe how others will be able to access and make use of your data, especially beyond the lifetime of the project itself. You should begin by explaining how you believe your data will be of value to others, how it will enhance your area of

---

[8] Advanced Computing Research Centre, https://www.acrc.bris.ac.uk/

research and what potential uses it may have in the future. Remember that it may not be necessary to share all the data that is produced during a project; considering whether it will be useful to others will assist you in deciding what should be made available.

Outline when you plan to make your data available. AHRC expect that data should be made available as soon as possible, but it is accepted there may be reasons why it cannot be accessible immediately, for example if you plan to publish research findings. Use this section of the DMP to justify any delays or embargoes. If a variety of audiences are expected to access your data, you should describe how these groups will be informed about its availability.

For all data which is to be shared, consider the costs involved and the expertise required to maintain and provide access beyond the end of the project (for example, issues relating to maintenance or updates to technology). There may also be a need to update the intellectual content (for example, information within a database) in addition to the technical infrastructure (for example, the format of a database).

Explain how each cost will be met and by whom. For example, the Library, or another University service, might provide a firm commitment to sustain and provide access to your data for a specified period.

The AHRC expects data to be made **freely available** (i.e. with open public access for data and open-source status for any software that you develop). If you propose to charge users for access, you will have to

justify this, as the default expectation is that all access will be free.

The easiest way to provide continued access to digital outputs is to publish in a data repository. Most can provide a Digital Object Identifier (DOI), a persistent identifier which allows digital outputs to be easily cited in publications. As of the 1st April 2013 all the UK's Research Funding Councils, as part of UKRI (then RCUK), require research outputs (i.e. journal articles) to provide a means by which third parties can access any underpinning research datasets. A Digital Object Identifier or DOI printed in a paper will lead an enquirer to a specific webpage where either the data is directly available, or that contains details of how the data can be accessed.

Given the extended timescales involved in the publication process (possibly extending beyond the mandatory three years mentioned above), it is strongly recommended that the authors of published academic outputs *do not provide their current contact details* as a means by which underpinning research data may be accessed, as these will change over time. It is also not recommended to use a project website as the only means of long-term access to your data, as these are rarely maintained for more than a few years beyond the lifetime of a project. If you're not planning to use an established data repository service to share your data, contact the Research Data Service[9] for further guidance.

---

[9] The University of Bristol's Research Data Service
data.bris.ac.uk

Archaeology Data Service

Award holders in archaeological disciplines should deposit outputs with the Archaeology Data Service (ADS). The ADS should be consulted before the start of the proposed research to discuss the form and extent of electronic materials to be deposited, as there will be a charge for this deposit.

Any significant archaeological resources or datasets funded by the AHRC (together with documentation) must be offered for deposit with the ADS within three months of the end of the project.

University of Bristol Research Data Repository

The University of Bristol has its own research data repository (data.bris) which researchers from any discipline may wish to use. This repository can provide ongoing access to research data for extended periods of time and can issue unique DOIs for deposited datasets. For smaller datasets, there is no cost. If you are planning to deposit larger datasets in the repository, a cost may be incurred. Contact the Research Data Service as early as possible if you believe you'll need to use Bristol's data repository.

As part of the process of depositing your research data and making it publicly accessible, the data.bris repository service automatically assigns a DOI (Digital Object Identifier) to your data, which can be used for citation purposes and for associating the dataset with other research outputs. Different access levels are available in the repository, from entirely open to

rigorously controlled, which is suitable for sensitive data.

Ethical and legal considerations

This section should highlight any ethical or legal issues around the collection, storage and sharing of data, including anonymity and following promises made to participants.

The AHRC will expect you, as award holder, to deal with any copyright issues that concern your research. Examples are ensuring that 'release forms' are signed for interviewees or ensuring that copyright permissions are provided with any video documentation that you produce as part of your project. Bear in mind that if you are planning to use existing data as part of your research, that data may be subject to copyright or other restrictions which could prevent you from sharing any new outputs you derive from it. You should inform the AHRC if this is the case.

Unless stated otherwise, the ownership of intellectual property lies with the organisation carrying out the research. However, if you plan to collaborate with an external partner, copyright and intellectual property rights issues may need to be clarified in a formal agreement. While this isn't required as part of your application, it should be noted that such an agreement will be created if the application is successful. Research Enterprise and Development[10] can advise further on collaborative research agreements and other intellectual property rights issues.

---

[10] Research Enterprise and Development, https://bristol.ac.uk/red/contracts

In addition, all recipients of research grants must adhere to the Data Protection Act 1998 (and, from May 2018, the General Data Protection Regulation). If you plan to handle sensitive, personal data, extra security measures must be considered. The Office of the University Secretary[11] can provide more advice on observing data protection legislation.

Including the correct information in consent forms is crucial for the potential sharing of data. Obtaining permission to publish data from human research participants is essential even if data is to be anonymised before publication. This is because some risk of re-identification may remain, even after anonymisation, and participants should be made aware that others outside of the research project may be able to view this data. Also, even if a participant has the right to withdraw from a study, it may not be possible to remove their data once it has been anonymised and amalgamated and in certain circumstances cannot then be excluded. The Research Data Service has produced a guide to sharing data involving human participants[12], which includes sample statements for consent forms.

Use your DMP to explain any anonymisation procedures that will take place prior to data being archived or shared. The exact information that needs to be removed during the process of anonymisation will vary depending on the contents of the dataset and the reason the unmodified data has been deemed unsuitable for sharing. The UK Data Service provide more guidance on effective anonymisation.[13]

Institutional expectations

When a DMP is submitted to the AHRC you will be asked to confirm that your institution has considered and will meet a number of requirements around research data management.[14] If these cannot be complied with, the proposal will be rejected. The requirements include: that the proposal is in line with an institutional data management policy; that data is stored appropriately with any risks considered; that legal, IP and ethical issues have been discussed with the relevant people at the institution; that you have consulted the institution's data support (i.e. Research Data Service). Unless the proposal is inherently digital in its methodology and naturally requires additional information around these points, you do not need to go into any further detail explaining these points in your DMP. The University of Bristol has the policies, infrastructure and services in place to meet all of these requirements.

[11] Office of the University Secretary, https://www.bristol.ac.uk/secretary/data-protection
[12] Sharing data concerning human participants, http://bit.ly/35hldfU

[13] UKDS Anonymisation https://www.ukdataservice.ac.uk/manage-data/legal-ethical/anonymisation
[14] AHRC Data Management Plan https://ahrc.ukri.org/documents/guides/data-management-plan/