

STFC Template

Plan Name STFC Template

Principal Investigator / Researcher TUoS Researcher

Funder -

Institution University of Sheffield

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.

DCC guidance on Data Type

Questions to consider:

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

Guidance:

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.

The University of Sheffield: guidance on Data Type

Please see the University of Sheffield webpage '[What is research data?](#)' for guidance.

Data preservation

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).

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?

Guidance:

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.

The University of Sheffield: guidance on Preservation Plan

Note: For guidance see the University of Sheffield webpages on '[Preserving your data](#)' and '[Data repositories](#)'.

Long term preservation and access may be best managed by using a specialist data repository. Some funder may specify a data repository to use, such as the [UK Data Service ReShare](#), [NERC Data Centres](#) or the [Archaeology Data Service](#). To find an appropriate repository, look in:

- [re3data.org](#)
- [BBSRC supported resources](#)
- [Wellcome Trust - Data repositories and database resources](#)

If no suitable repository is available you may [deposit data in ORDA](#), the University of Sheffield data repository. Alternatively, if you need to regulate users' access through 'Data sharing agreements', data may be retained in the University's research storage infrastructure and [registered in ORDA](#).

Suggested text in all cases: *"Data will be archived in line with the University of Sheffield's Research Data Management Policy, which is a component of the University's Policy on Good R&I Practices (the 'GRIP' Policy)."*

Where data is in paper format: *"Data collected in paper form will be routinely digitised and the paper form disposed of / stored for at least 10 years at our universities in secured areas."*

For data deposited in external data repositories: *"Research data selected for long-term preservation and sharing will be deposited in [name of repository/weblink]. The [name of repository] is openly accessible and searchable and will guarantee preservation of these data for ten years or more. Metadata records describing these data will be created in ORDA, the University of Sheffield research data registry and repository"*

Where some research data are being deposited in ORDA: *"Data that are not*

deposited in [name of repository/weblink] will be deposited in ORDA, a repository and registry of research data produced at the University of Sheffield, which will preserve data for ten years or more.”

Where data is deposited in ORDA only: *“Data selected for long-term preservation and sharing will be deposited in ORDA, a repository and registry of research data produced at the University of Sheffield, which will guarantee preservation for ten years or more.”*

Where data is being retained locally, but not made ‘openly’ accessible: *“Data selected for long-term preservation and sharing will be stored on centrally provisioned University of Sheffield virtual servers and research storage infrastructure (<https://www.sheffield.ac.uk/cics/research>) for at least ten years. Records of these data will be published in ORDA, a registry of research data produced at the University of Sheffield.”*

DCC guidance on Data Repository

Questions to consider:

- Where (i.e. in which repository) will the data be deposited?

Guidance:

Most research funders recommend the use of established data repositories, community databases and related initiatives to aid data preservation, sharing and reuse.

An international list of data repositories is available via [Databib](#) or [Re3data](#).

The University of Sheffield: guidance on Data Repository

Note: For guidance see the University of Sheffield webpages on '[Publishing and sharing your research data](#)' and '[Data repositories](#)'.

Long term preservation and access may be best managed by using a specialist data repository. Some funder may specify a data repository to use, such as the [UK Data Service ReShare](#), [NERC Data Centres](#) or the [Archaeology Data Service](#). To find an appropriate repository, look in:

- re3data.org
- [BBSRC supported resources](#)
- [Wellcome Trust - Data repositories and database resources](#)

If no suitable repository is available you may [deposit data in ORDA](#), the University of Sheffield data repository. Alternatively, if you need to regulate users' access through 'Data sharing agreements', data may be retained in the University's research storage infrastructure and [registered in ORDA](#).

DCC guidance on Data Selection

Questions to consider:

- Which data are of long-term value and should be shared and/or preserved?
- How will you decide what to keep?

Guidance:

Indicate which data you intend to preserve beyond the period of funding. This should be based on what has long-term value and is economically viable to keep. Consider how long you wish to keep the data and what will happen to it e.g. deposit in a data repository to enable reuse.

See the DCC guide: [How to appraise and select research data for curation](#).

The University of Sheffield: guidance on Data Selection

Most funders now expect data underlying published papers, plus data of particular long-term value, to be made available to other researchers at the end of a project. Please see the University of Sheffield webpage on '[Preserving your data](#)' for guidance on data selection. Check your funder's policy on the University of Sheffield webpage '[Research funder policy summaries](#)'.

The Digital Curation Centre provides useful advice about [data selection and appraisal](#). The [NERC data value checklist](#) provides guidance on determining long-term value.

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.

DCC guidance on Metadata

Questions to consider:

- How will you capture / create the metadata?
- Can any of this information be created automatically?
- What metadata standards will you use and why?

Guidance:

Metadata should be created to describe the data and aid discovery. Consider how you will capture this information and where it will be recorded e.g. in a database with links to each item, in a 'readme' text file, in file headers etc.

Researchers are strongly encouraged to use community standards to describe and structure data, where these are in place. The DCC offers a [catalogue of disciplinary metadata standards](#).

The University of Sheffield: guidance on Metadata

Metadata is a structured form of documentation that identifies and describes your data.

Researchers should use community standards, where they exist: see the DCC webpage on [Disciplinary metadata standards](#).

Please see the University of Sheffield webpages on '[Describing your data](#)' and '[Metadata](#)' for guidance.

DCC guidance on Documentation

Questions to consider:

- What metadata, documentation or other supporting material should accompany the data for it to be interpreted correctly?
- What information needs to be retained to enable the data to be read and interpreted in the future?

Guidance:

Describe the types of documentation that will accompany the data to provide secondary users with any necessary details to prevent misuse, misinterpretation or confusion. This may include information on the methodology used to collect the data, analytical and procedural information, definitions of variables, units of measurement, any assumptions made, the format and file type of the data.

The University of Sheffield: guidance on Documentation

Note: Documentation and metadata describe the context, content and structure of your data and are essential for understanding and reusing them. See the University of Sheffield webpage '[Describing your data](#)' for more information.

Example text: *"Methods and SOPs will be stored electronically in Microsoft Word documents (.doc) with the spreadsheets containing data"*

"Explanation of the experimental and analytical methods used will be provided in text documents, stored alongside the data"

"Data documentation will accompany datasets submitted to the ... repository at the end of the research"

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.

DCC guidance on Period of Preservation

Questions to consider:

- How long will the data be retained and preserved?

Guidance:

This may depend on the type of data. Most research funders expect data 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.

The University of Sheffield: guidance on Period of Preservation

The [RCUK](#) funders generally expect data that underpins findings in publications should be accessible for at least ten years after publication. However, data that by their nature cannot be re-measured or recreated such as earth observations or people-based data may often warrant indefinite storage and preservation.

Many research funders specify which data need preserving, how long for and where to deposit these data: See the University of Sheffield webpage on '[Research funder policy summaries](#)' for information.

Data sharing

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.

DCC guidance on Expected Reuse

Questions to consider:

- Who may be interested in using your data?
- What are the further intended or foreseeable research uses for the data?

Guidance:

You should think about the possibilities for reuse of your data in other contexts and by other users, and connect this as appropriate with your plans for dissemination and Pathways to Impact. Where there is potential for reuse, you should use standards and formats that facilitate this. Where possible outline the types of users you expect and estimate numbers.

The University of Sheffield: guidance on Expected Reuse

The potential of research data for reuse must be considered when planning long-

term curation. The Digital Curation Centre provides useful advice about [data selection and appraisal](#). The [NERC data value checklist](#) provides guidance on determining long-term value.

DCC guidance on Data Selection

Questions to consider:

- Which data are of long-term value and should be shared and/or preserved?
- How will you decide what to keep?

Guidance:

Indicate which data you intend to preserve beyond the period of funding. This should be based on what has long-term value and is economically viable to keep. Consider how long you wish to keep the data and what will happen to it e.g. deposit in a data repository to enable reuse.

See the DCC guide: [How to appraise and select research data for curation](#).

The University of Sheffield: guidance on Data Selection

Most funders now expect data underlying published papers, plus data of particular long-term value, to be made available to other researchers at the end of a project. Please see the University of Sheffield webpage on '[Preserving your data](#)' for guidance on data selection. Check your funder's policy on the University of Sheffield webpage '[Research funder policy summaries](#)'.

The Digital Curation Centre provides useful advice about [data selection and appraisal](#). The [NERC data value checklist](#) provides guidance on determining long-term value.

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.

DCC guidance on Timeframe For Data Sharing

Questions to consider:

- When will you make the data available?

Guidance:

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.

The University of Sheffield: guidance on Timeframe For Data Sharing

Note: At the end of your research project, your funder may require you to make your research data available for sharing with as few restrictions as possible. 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 University of Sheffield '[Research funder policy summaries](#)' webpage to determine when you need to make your data available.

Suggested text in all cases: *“The project group (including collaborators) will have exclusive use of the data until the main research findings are published or patent applications have been filed [if potentially relevant to project]” and/or “...or for a period of x months/years.”*

Suggested text if delays are foreseen: *“Delays in sharing data may arise through a delayed ability to analyse or publish the research findings.” and/or “Delays in sharing data may arise due to IPR and if this is a factor, advice will be sought from the University’s Research & Innovation Services.”*

Optional additional text: *“Following publication, data will be made available on request or shared through the [relevant research platforms].”*

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.

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?

Guidance:

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.

The University of Sheffield: guidance on Method For Data Sharing

Note: At the end of your research project, your funder may require you to make your research data available for sharing with as few restrictions as possible. Data may be shared by being published in:-

- a Repository or Data Centre - see the University of Sheffield webpage on '[Research data repositories](#)' for guidance
- a journal as an article's supplementary material
- a data journal as a data paper.

Wherever data is published, a metadata record should be [registered in ORDA](#), the University of Sheffield data repository.

Suggested text for use when data will be placed in a repository: *“Data will be made available through shared research platforms [insert repository / platform relevant to project] with the relevant permissions in place.”*

Suggested text for use when data will not be placed in a repository: *“The lead PI and project team [including collaborators if applicable] will review applications to access experimental data and make the decision on whether to supply research data to potential applicants. Data will then be released on a case by case basis.”*

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.

DCC guidance on Resourcing

Questions to consider:

- What additional resources are needed to deliver your plan?
- Is additional specialist expertise (or training for existing staff) required?
- Do you have sufficient storage and equipment or do you need to cost in more?
- Will charges be applied by data repositories?
- Have you costed in time and effort to prepare the data for sharing / preservation?

Guidance:

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 that may be charged by IT services. 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](#).

The University of Sheffield: guidance on Resourcing

The University of Sheffield [research data storage facility](#) allocates 10TB storage free to research groups during the lifetime of a project. If a larger quota is required then this will involve charges. Long-term archiving of data may involve charges also. Get in touch with CiCS to discuss your requirements and get a quote at <https://www.sheffield.ac.uk/cics/support/help>.

[ORDA](#), the University of Sheffield research data repository is free to use. You should enquire about charges made by other data repositories you intend to use.