Horizon 2020 DMP

Plan Name Horizon 2020 DMP - Horizon 2020 DMP

Principal Investigator / **Researcher** TUoS Researcher

Funder -

Institution University of Sheffield

1. Data summary

State the purpose of the data collection/generation

purpose of the data collection/generation

Explain the relation to the objectives of the project

relation to the objectives of the project

Specify the types and formats of data generated/collected

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.

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?

Guidance:

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.

The University of Sheffield: guidance on Data Format

Please see the University of Sheffield webpage on 'Organising your data: Choosing data formats' for guidance.

Specify if existing data is being re-used (if any)

DCC guidance on Existing Data

Questions to consider:

- Are there any existing data or methods that you can reuse?
- Do you need to pay to reuse existing data?
- Are there any restrictions on the reuse of third-party data?
- Can the data that you create - which may be derived from third-party data - be shared?

Guidance:

Check to see if there are any existing data that you can reuse, for examples by consulting relevant repositories. When creating new data sources, explain why existing data sources cannot be reused. If purchasing or reusing existing data sources, explain how issues such as copyright and IPR have been addressed. A list of repositories is provided by Databib or Re3data.

The University of Sheffield: guidance on Existing Data

Please see the University of Sheffield webpage on 'Finding and reusing data' for guidance.

Specify the origin of the data

origin of the data

State the expected size of the data (if known)

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?

Guidance:

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.

The University of Sheffield: guidance on Data Volumes

Please see the University of Sheffield Corporate Information and Computing Services webpages on 'Research data storage' and 'Storage options' for guidance.

Outline the data utility: to whom will it be useful

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.

2.1 Making data findable, including provisions for metadata [FAIR data]

Outline the discoverability of data (metadata provision)

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.

Outline the identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers?

Identifiability of data, standard identification mechanism, persistent and unique identifiers and Digital Object Identifiers.

Outline naming conventions used

Naming conventions used

Outline the approach towards search keyword

DCC guidance on Discovery by Users

Questions to consider:

- How will potential users find out about your data?
- Will you provide metadata online to aid discovery and reuse?

Guidance:

Indicate how potential new users can find out about your data and identify whether they could be suitable for their research purposes. For example, you may provide basic discovery metadata online (i.e. the title, author, subjects, keywords and publisher).

The University of Sheffield: guidance on Discovery by Users

Note: All research data selected for long-term preservation should be registered in the University of Sheffield research data repository ORDA. A metadata record should be created in ORDA irrespective of whether the data files are deposited in ORDA or in another repository. Research data in non-digital formats and digital data that cannot be made accessible or requires controlled access should also be registered in ORDA. Please see the University of Sheffield webpage on 'ORDA user guidance'.

Suggested text in all cases: *"Records of datasets will be published in ORDA, the University of Sheffield's registry of research data produced at the University, which will issue DataCite DOIs for registered datasets and promote discovery."*

Outline the approach for clear versioning

Approach for clear versioning

Specify standards for metadata creation (if any). If there are no standards in your discipline describe what metadata will be created and how

EC Guidance

The Research Data Alliance provides a Metadata Standards Directory that can be searched for discipline-specific standards and associated tools.

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.

2.2 Making data openly accessible [FAIR data]

Specify which data will be made openly available? If some data is kept closed provide rationale for doing so

EC Guidance

Participating in the ORD Pilot does not necessarily mean opening up all your research data. Rather, the ORD pilot follows the principle "**as open as possible, as closed as necessary**" and focuses on encouraging sound data management as an essential part of research best practice.

The Commission recognises that there are good reasons to keep some or even all research data generated in a project closed. Where data need to be shared under restrictions, explain why, clearly separating legal and contractual reasons from voluntary restrictions.

Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if relevant provisions are made in the consortium agreement and are in line with the reasons for opting out.

Specify how the data will be made available

EC Guidance

For example by deposition in a repository. The Registry of Research Data Repositories provides a useful listing of repositories that you can search to find a place of deposit.

If you plan to deposit in a repository, it is useful to explore appropriate arrangements with the identified repository in advance.

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."*

Specify what methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)?

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 webapge '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 where the data and associated metadata, documentation and code are deposited

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 funders specify a data repository to use, such as UK Data Service ReShare, NERC Data Centres or 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.

Specify how access will be provided in case there are any restrictions

EC Guidance
For example is there a need for a data access committee.

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?

Guidance:

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.

The University of Sheffield: guidance on Managed Access Procedures

Note: Restrictions on the release of data may be allowed, to protect confidentiality and for other ethical and legal considerations. Access to and use of Sensitive and confidential data can be restricted and regulated using end user licenses, data sharing agreements or by using a data enclaves.

Whatever form of publishing is used, research data needs to be licensed to indicate what users may or may not do with the data. Data repositories will indicate what licenses are available for the data they house. More information is available from the Digital Curation Centre webpage 'How to license research data'.

See the University of Sheffield 'Regulatory requirements' webpage and the UK Data Service 'Legal and ethical issues' webpages for more information.

Suggested text: *"The University of Sheffield's Good Research and Innovation Practice (GRIP) Policy follows RCUK principles for data sharing (http://www.rcuk.ac.uk/research/datapolicy/)"*

Suggested text where a Data sharing agreement is required: *"External users will be bound by data sharing agreements as specified by the [name of funder] Data Sharing Policy."*

Suggested text where an external collaborator is involved: *"Data sharing agreements will be put in place with [name of collaborator], who will be a primary re-user of data"*

Faculty of Medicine, Dentistry & Health: guidance on Managed Access Procedures

Access to and use of Sensitive and confidential data can be restricted and regulated using end user licenses or data sharing agreements.

MRC, Wellcome Trust and Cancer Research UK provide guidance on data sharing agreements in 'Good Practice Principles for Sharing Individual Participant Data from

Publicly Funded Clinical Trials'.

MRC provides information on data access governance, facilitating data access and data sharing agreements in 'MRC Policy and Guidance on Sharing of Research Data from Population and Patient Studies' and 'Principles for access to, and use of, MRC funded research data'.

The University of Sheffield ScHARR Information Governance Policy provides guidance on information sharing and an example data sharing agreement.

2.3 Making data interoperable [FAIR data]

Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability.

EC Guidance

Inteoperability means allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins.

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.

Specify whether you will be using standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?

Using standard vocabulary for all data types present in data set, to allow inter-disciplinary interoperability / Providing mapping to more commonly used ontologies.

2.4 Increase data re-use (through clarifying licenses) [FAIR data]

Specify how the data will be licenced to permit the widest reuse possible

EC Guidance

The EUDAT B2SHARE tool includes a built-in license wizard that facilitates the selection of an adequate license for research data.

DCC guidance on IPR Ownership and Licencing

Questions to consider:

- Who owns the data?
- How will the data be licensed for reuse?
- If you are using third-party data, how do the permissions you have been granted affect licensing?
- Will data sharing be postponed / restricted e.g. to seek patents?

Guidance:

State who will own the copyright and IPR of any new data that you will generate. For multi-partner projects, IPR ownership may be worth covering in a consortium agreement. If purchasing or reusing existing data sources, consider how the permissions granted to you affect licensing decisions. Outline any restrictions needed on data sharing e.g. to protect proprietary or patentable data.

See the DCC guide: How to license research data.

The University of Sheffield: guidance on IPR Ownership and Licencing

Please see the University of Sheffield webpage on 'Regulatory requirements' for guidance on IPR and data licensing.

The UK Data Service provides guidance on 'Rights relating to research data' including licensing.

The Digital Curation Centre provides guidance on 'How to license research data'.

Specify when the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed

EC Guidance

Reasons for embargoes may include time to publish or seek patents. If an embargo is sought, specify why and for how long, bearing in mind that research data should be made available as soon as possible.

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 whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why

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?

Guidance:

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.

The University of Sheffield: guidance on Restrictions on 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. Restrictions on the release of data may be allowed, to protect confidentiality and for other ethical and legal considerations:-

- Does your data include confidential and sensitive information?
- Have participants given consent for their data being shared?
- Consider what can be done to make sensitive data openly sharable - can these data be anonymised?
- If different parts of your research data require different access conditions, separate them and deposit them separately, applying different access conditions.

See the University of Sheffield 'Regulatory requirements' webpage and the UK Data Service 'Legal and ethical issues' webpages for more information.

Suggested text if no restrictions are foreseen: *"At present we do not foresee any delays in data sharing following publication of the main research findings."*

Suggested text for patient-based studies: *"Patients will be made aware of our data sharing procedures at the time of consent."*

Describe data quality assurance processes

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?

Guidance:

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.

The University of Sheffield: guidance on Data Quality

Please see the UK Data Service webpage on data quality assurance for guidance.

Specify the length of time for which the data will remain re-usable

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.

3. Allocation of resources

Estimate the costs for making your data FAIR. Describe how you intend to cover these costs

EC Guidance

Note that costs related to open access to research data are eligible as part of the Horizon 2020 grant (if compliant with the Grant Agreement conditions).

Costs are eligible for reimbursement during the duration of the project under the conditions defined in the H2020 Grant Agreement, in particular Article 6 and Article 6.2.D.3, but also other articles relevant for the cost category chosen.

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.

Clearly identify responsibilities for data management in your project

DCC guidance on Responsibilities

Questions to consider:

- Who is responsible for each data management activity?
- How are responsibilities split across partner sites in collaborative research projects?

Guidance:

Outline the roles and responsibilities for all activities e.g. data capture, metadata production, data quality, storage and backup, data archiving & data sharing. Individuals should be named where possible. For collaborative projects you should explain the co-ordination of data management responsibilities across partners.

See UKDS guidance on data management roles and responsibilities.

Describe costs and potential value of long term preservation

EC Guidance

Consider who decides what data will be kept and for how long

4. Data security

Address data recovery as well as secure storage and transfer of sensitive data

EC Guidance

Also consider whether the data is safely stored in certified repositories for long term preservation and curation.

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?

Guidance:

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.

See UKDS guidance on data security.

The University of Sheffield: guidance on Data Security

Note: Data security is needed to prevent unauthorised access or disclosure and changes to or destruction of data. Please see the University of Sheffield webpage 'Keeping your data safe' for guidance.

All staff and researchers must complete the online training at https://infosecurity.shef.ac.uk. This training will be particularly helpful when filling this part of the DMP. If you require assistance please contact the University's Information Security team

at https://www.shef.ac.uk/cics/infosec.

The University has policies relating to information security requiring its users to adhere, as a minimum, to the following security standards: Information Security Policy, Data Protection Policy. More secure system policies may be defined where necessary, for example where patient data is involved. University Departments may have established their own policies regarding information security, e.g. ScHARR Information Governance Policy.

The University of Sheffield is not an accredited ISO 27001 institution. The University provides email, Contacts and calendaring services, Google Drive and Google Sites through the Google Apps for Education suite. Google Apps for Education (and the data centres that support the service) are SSAE 16 / ISAE 3402 Type II SOC 2 audited and have achieved ISO 27001 certification. The University is satisfied that personal data is being processed appropriately in accordance with UK Data Protection Law and the University's own privacy policies, and that the security controls put in place by Google are sufficient to protect University data.

Example text for high-risk data: *"We will comply with the Data Protection Act and the University's own Information Security and Data Protection Policies. The project is governed by Department of Health and so we will comply with the NHS IGT."*

Example text for high-risk data: *"We recognize that this data is highly confidential and is critical to the clinical treatment of patients. Therefore a project specific security policy has been developed in conjunction with the University's Information Security Team [link to policy]".*

Suggested text for low-risk data: *"The data will will not include personal data relating to human participants. The University's Information Security Policies will be abided by at all times."*

Faculty of Medicine, Dentistry & Health: guidance on Data Security

Some Faculty of Medicine, Dentistry and Health Departments have established their own policies regarding information security, e.g. ScHARR Information Governance Policy.

DCC guidance on Storage and Backup

Questions to consider:

- Where will the data be stored?
- How will the data be backed up? i.e. how often, to where, how many copies, is this automated…
- Who will be responsible for storage and backup?
- Do you have access to enough storage or will you need to include charges for additional services?

Guidance:

Describe how the data will be stored and backed-up to ensure the data and metadata are securely stored during the lifetime of the project. Storing data on laptops, computer hard drives or external storage devices alone is very risky. The use of robust, managed storage with automatic backup, for example that provided by university IT teams, is preferable.

See UKDA guidance on data storage and backup.

The University of Sheffield: guidance on Storage and Backup

Note: Storing data on laptops, computer hard drives or external storage devices alone is not recommended. The use of robust, managed storage with automatic backup is preferred by the University and by funders.

Data and definitive project documentation should be stored on centrally provisioned University of Sheffield virtual servers and research data storage infrastructure throughout the lifetime of the project. Both Windows and Linux Virtual Servers with up to 10TB of storage are made available to research projects. Access control is by authorised University computer account username and password. Off-site access is facilitated by secure VPN connection authenticated by University username and remote password. By default, two copies of data are kept across two physical plant rooms, with a 28 day snapshot made of data and backed up securely offsite at least daily. This service is maintained by the University's Corporate Information and Computing Services.

Google Drive may be used for more flexible collaborative working but only where non personal-sensitive information is involved. Where Google Drive is used, copies of complete and definitive documents should be transferred to the main project repository on the University research storage infrastructure.

Please see the University of Sheffield webpage on 'Keeping your data safe' for further guidance.

Faculty of Medicine, Dentistry & Health: guidance on Storage and Backup

All requests for research data storage in the Faculty of Medicine, Dentistry and Health should be made to the Faculty IT Hub in the first instance (med-it@sheffield.ac.uk). They will work with you to create an appropriate folder structure and give access to authorised users.

5. Ethical aspects

To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former

EC Guidance

Consider whether there are any ethical or legal issues than can have an impact on data sharing. For example, is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data?

DCC guidance on Ethical Issues

Questions to consider:

- Have you gained consent for data preservation and sharing?
- How will sensitive data be handled to ensure it is stored and transferred securely?
- How will you protect the identity of participants? e.g. via anonymisation or using managed access procedures

Guidance:

Investigators carrying out research involving human participants must ensure that consent is obtained to share data. Managing ethical concerns may include: anonymisation of data; referral to departmental or institutional ethics committees; and formal consent agreements. Ethical issues may affect how you store data, who can see/use it and how long it is kept. You should show that you're aware of this and have planned accordingly.

See UKDS guidance on consent for data sharing

The University of Sheffield: guidance on Ethical Issues

For guidance on ethical and legal issues please see the University of Sheffield webpage on 'Regulatory requirements'.

See also the UK Data Service webpages on 'Legal and ethical issues'.

6. Other

Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any)

DCC guidance on Related Policies

Questions to consider:

- Are there any existing procedures that you will base your approach on?
- Does your department/group have data management guidelines?
- Does your institution have a data protection or security policy that you will follow?
- Does your institution have a Research Data Management (RDM) policy?
- Does your funder have a Research Data Management policy?
- Are there any formal standards that you will adopt?

Guidance:

List any other relevant funder, institutional, departmental or group policies on data management, data sharing and data security. Some of the information you give in the remainder of the DMP will be determined by the content of other policies. If so, point/link to them here.

The University of Sheffield: guidance on Related Policies

Note: Please refer to the University of Sheffield policies below: these may be used when writing a DMP for any funder. For the 'Data sharing policy', add the data sharing policy of your funder. Please see the University of Sheffield webpage on 'Expectations and obligations' for information on funder policies. List departmental or research group policies on data management, sharing and security here also.

Data Management Policy & Procedures:

University of Sheffield Research Data Management Policy http://www.shef.ac.uk/polopoly_fs/1.553350!/file/GRIPPolicyextractRDM.pdf

University of Sheffield Research Data Management Guidance http://www.sheffield.ac.uk/library/rdm/index

Data Security Policies:

University of Sheffield Data protection policy http://www.shef.ac.uk/cics/dataprotection

University of Sheffield Information Security
policy http://www.shef.ac.uk/cics/policies/infosecpolicy

Data Sharing Policy:

The study will adhere to the RCUK principles http://www.rcuk.ac.uk/research/datapolicy/

Institutional Information Policy:

University of Sheffield Good Research and Innovation Practice (GRIP)
Policy http://www.sheffield.ac.uk/polopoly_fs/1.356709!/file/GRIPPolicySenateapproved.pdf
University of Sheffield Information Security
Policies http://www.shef.ac.uk/cics/policies/infosec

Institutional Ethics Policy:

The University of Sheffield Ethics Policy Governing Research Involving Human Participants, Personal Data and Human Tissue http://www.shef.ac.uk/ris/other/gov-ethics/ethicspolicy

Faculty of Medicine, Dentistry & Health: guidance on Related Policies

ScHARR Information Governance
Policy https://www.shef.ac.uk/scharr/research/igov/policy00

NHS Information Governance Toolkit https://www.igt.hscic.gov.uk/