Digital Curation Centre: DCC Template

Data Collection

What data will you collect or create?

Guidance:

Questions to consider:

- What type, format and volume of data?
- Do your chosen formats and software enable sharing and long-term access to the data?
- Are there any existing data that you can reuse?

Guidance:

Give a brief description of the data, including any existing data or third-party sources that will be used, in each case noting its content, type and coverage. Outline and justify your choice of format and consider the implications of data format and data volumes in terms of storage, backup and access.

How will the data be collected or created?

Guidance:

Questions to consider:

- What standards or methodologies will you use?
- How will you structure and name your folders and files?
- How will you handle versioning?
- What quality assurance processes will you adopt?

Guidance:

Outline how the data will be collected/created and which community data standards (if any) will be used. Consider how the data will be organised during the project, mentioning for example naming conventions, version control and folder structures. 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.

Documentation and Metadata

What documentation and metadata will accompany the data?

Guidance:

Questions to consider:

- What information is needed for the data to be to be read and interpreted in the future?
- How will you capture / create this documentation and metadata?
- What metadata standards will you use and why?

Guidance:

Describe the types of documentation that will accompany the data to help secondary users to understand and reuse it. This should at least include basic details that will help people to find the data, including who created or contributed to the data, its title, date of creation and under what conditions it can be accessed.

Documentation may also include details on the methodology used, analytical and procedural information, definitions of variables, vocabularies, units of measurement, any assumptions made, and the format and file type of the data. Consider how you will capture this information and where it will be recorded. Wherever possible you should identify and use existing community standards.

Ethics and Legal Compliance

How will you manage any ethical issues?

Guidance:

Questions to consider:

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

Guidance:

Ethical issues affect how you store data, who can see/use it and how long it is kept. Managing ethical concerns may include: anonymisation of data; referral to departmental or institutional ethics committees; and formal consent agreements. You should show that you are aware of any issues and have planned accordingly. If you are carrying out research involving human participants, you must also ensure that consent is requested to allow data to be shared and reused.

How will you manage copyright and Intellectual Property Rights (IPR) issues?

Guidance:

Questions to consider:

- Who owns the data?
- How will the data be licensed for reuse?
- Are there any restrictions on the reuse of third-party data?
- Will data sharing be postponed / restricted e.g. to publish or seek patents?

Guidance:

State who will own the copyright and IPR of any data that you will collect or create, along with the licence(s) for its use and reuse. For multi-partner projects, IPR ownership may be worth covering in a consortium agreement. Consider any relevant funder, institutional, departmental or group policies on copyright or IPR. Also consider permissions to reuse third-party data and any restrictions needed on data sharing.

Storage and Backup

How will the data be stored and backed up during the research?

Guidance:

Questions to consider:

- Do you have sufficient storage or will you need to include charges for additional services?
- How will the data be backed up?
- Who will be responsible for backup and recovery?
- How will the data be recovered in the event of an incident?

Guidance:

State how often the data will be backed up and to which locations. How many copies are being made? Storing data on laptops, computer hard drives or external storage devices alone is very risky. The use of robust, managed storage provided by university IT teams is preferable. Similarly, it is normally better to use automatic backup services provided by IT Services than rely on manual processes. If you choose to use a third-party service, you should ensure that this does not conflict with any funder, institutional, departmental or group policies, for example in terms of the legal jurisdiction in which data are held or the protection of sensitive data.

How will you manage access and security?

Guidance:

Questions to consider:

- What are the risks to data security and how will these be managed?
- How will you control access to keep the data secure?
- How will you ensure that collaborators can access your data securely?
- If creating or collecting data in the field how will you ensure its safe transfer into your main secured systems?

Guidance:

If your data is confidential (e.g. personal data not already in the public domain, confidential

information or trade secrets), you should outline any appropriate security measures and note any formal standards that you will comply with e.g. ISO 27001."

Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

Guidance:

Questions to consider:

- What data must be retained/destroyed for contractual, legal, or regulatory purposes?
- How will you decide what other data to keep?
- What are the foreseeable research uses for the data?
- How long will the data be retained and preserved?

Guidance:

Consider how the data may be reused e.g. to validate your research findings, conduct new studies, or for teaching. Decide which data to keep and for how long. This could be based on any obligations to retain certain data, the potential reuse value, what is economically viable to keep, and any additional effort required to prepare the data for data sharing and preservation. Remember to consider any additional effort required to prepare the data for sharing and preservation, such as changing file formats.

What is the long-term preservation plan for the dataset?

Guidance:

Questions to consider:

- Where e.g. in which repository or archive will the data be held?
- What costs if any will your selected data repository or archive charge?
- Have you costed in time and effort to prepare the data for sharing / preservation?

Guidance:

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.

Data Sharing

How will you share the data?

Guidance:

Questions to consider:

- How will potential users find out about your data?
- With whom will you share the data, and under what conditions?
- Will you share data via a repository, handle requests directly or use another mechanism?
- When will you make the data available?
- Will you pursue getting a persistent identifier for your data?

Guidance:

Consider where, how, and to whom data with acknowledged long-term value should be made available. The methods used to share data will be dependent on a number of factors such as the type, size, complexity and sensitivity of data. If possible, mention earlier examples to show a track record of effective data sharing. Consider how people might acknowledge the reuse of your data.

Are any restrictions on data sharing required?

Guidance:

Questions to consider:

- What action will you take to overcome or minimise restrictions?

- For how long do you need exclusive use of the data and why?
- Will a data sharing agreement (or equivalent) be required?

Guidance:

Outline any expected difficulties in sharing data with acknowledged long-term value, along with causes and possible measures to overcome these. Restrictions may be due to confidentiality, lack of consent agreements or IPR, for example. Consider whether a non-disclosure agreement would give sufficient protection for confidential data.

Responsibilities and Resources

Who will be responsible for data management?

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?

Guidance:

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

What resources will you require to deliver your plan?

Guidance:

Questions to consider:

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

Guidance:

Carefully consider any resources needed to deliver the plan, e.g. software, hardware, technical expertise, etc. Where dedicated resources are needed, these should be outlined and justified.