

# DATUM DATA MANAGEMENT PLAN (DMP) GUIDANCE NOTES FOR COMPLETION

## Purpose of a Data Management Plan

The purpose of a DMP is:

- to assist in planning the research data management (RDM) aspects of your research
- to assist you in making RDM decisions
- to identify the RDM actions required
- to highlight areas that need further thought
- to provide a record of decisions made and actions taken

The DATUM in Action DMP is based on the Digital Curation Centre's template.<sup>i</sup>

## Completing the DMP

The sections in this guidance correspond to the sections within the DMP template.

General notes for completing the DMP:

1. This is a generic template to cover all types of projects/disciplines. The first section of the DMP 'Applicable Sections Within the DMP' enables you to see all the sections of the DMP at a glance. If a section is not relevant to your project mark it as not applicable.
2. In your first draft you may not know the answer to all the questions. These will be: (i) issues for you to discuss with your project team/supervisory team to obtain agreement on the answers; or (ii) will become clear as the project develops. A DMP is a living document.
3. Instead of duplicating information in the DMP that exists in other documents/records, we suggest you just give the location of these documents where this information can be found: (i) For documents on a computer, give the name of the computer (e.g. shared drive) and the folder name and path: give the highest level folder for that topic (e.g. R:/IMI Research Group/DATUM in action/Research/DMP); (ii) For paper documents, give the physical location (e.g. filing cabinet in a named office); (iii) For Web items, give the URL.
4. Examples have been provided (highlighted in grey) to provide some guidance within the DMP, in addition to this more detailed guidance. And in the action boxes, suggested actions have been given to make completing the DMP easier.

### Applicable Sections Within the DMP

This table lists all the sections within the DMP template, to enable you to decide which sections are applicable to your research.

Some sections are applicable to all projects, and these have already been marked with a 'Yes'.

## 1. Introduction and Context

### 1.1 Project Information

This section is applicable to all projects.

Most of the sub-sections are self explanatory. Brief notes are given here as follows:

- Name of organisation funding project

Many bodies could fund research, e.g. a Research Council, a KTP (Knowledge Transfer Partnership), a charity, a private company, internal University sources. Research could also not have any funding but be carried out as a self-funded activity or for RSA (Research and Scholarly activity leave): in that case just note this in this sub-section.

- Brief outline of project's aims and objectives

These can be found in project documents so it is not necessary to include them here unless there is a requirement for you to do so (e.g. in a funder's DMP requirements).

- Location of project documents that provide the above type of information

Such documents could include: proposal, project plan, project summary, research team contact details etc.

### 1.2 Data Management Plan (DMP) Information

#### 1.2.1 Requirements for a DMP

This section may, or may not, be applicable to your project.

There might be a requirement to produce a plan, or to have some general mechanism for data management, e.g. from the University or a funding body. This section only refers to the requirement to produce a DMP. There may be other data management requirements from such organisations but these will be covered in other sections of this DMP.

#### 1.2.2 Allocation of Research Data Management (RDM) Roles and Responsibilities

This section is applicable to all projects.

The Principal Investigator (PI) (or sole researcher where applicable) is ultimately **accountable** for the RDM of the project. However, they can delegate RDM roles (and the responsibility for undertaking those roles/tasks) to other project team members., e.g. one person may be given the responsibility for completing the DMP, in consultation with other project team members, and keeping it up to date.

An example of RDM roles and responsibilities are:

#### Principal Investigator

The PI's role and responsibilities are to:

- Have overall accountability for research data management (RDM) and ensure that it is implemented
- Set the RDM culture and practice (based on good practice guidance) and delegate responsibilities as appropriate
- Make decisions (after consultation within the team) about issues such as data access, data sharing, long-term retention of data

- Develop (in consultation with the team) new RDM protocols/procedures and/or use/amend existing protocols: a team member could be delegated to scope and write these documents
- Ensure a DMP is developed and maintained: a team member could be delegated to do this
- Ensure that all RDM requirements (legal, funder, institutional, etc.) are included in the DMP
- Monitor RDM practices and ensure RDM requirements are met
- Manage and/or devolve responsibility for the project level documents and records i.e.
  - proposal, action plan etc.
  - reporting
  - financial
  - presentations, publications, outputs
- Obtain the necessary resources to conduct RDM, e.g. IT, allocation of responsibility for a team member to be a data controller

## Researchers

Researchers' responsibilities are to:

- Follow the University policy and recommended best practice on RDM on an individual basis
- Follow good RDM practice on an individual basis
- Follow the project's RDM practices for project activities (e.g. collecting/creating data, analysing data, producing outputs)
- Raise RDM issues with the PI for resolution
- Undertake RDM roles as delegated by the PI

### 1.2.3 DMP Version Tracking

This section is applicable to all projects.

As a living document, the DMP will change throughout the lifetime of the project. Therefore, the different version of the DMP will need to be tracked. Recording date information and creating different copies of the DMP with version numbers in the file name is important for version control and placing the DMP in context.

Be proportionate. How much version information do you need to record here in this section? Do you need to keep earlier versions? We suggest that there are 3 important points in a project where a copy of the then current DMP should be kept and formally signed off, i.e.:

- Completion of project proposal
- Completion of planning stage of project
- End of project

See the end of the DMP for the 'Signing off' Section.

## 2. Ethical and Legal Issues

### 2.1 Ethical and Legal Risk Factors

This section is applicable to all projects.

This table enables you to see at a glance the main ethical or legal risk factors that could be relevant to your project (and then directs you to the appropriate section where the issue is covered). Mark which ones are applicable/not applicable.

## 2.2 Ethical Issues Related to Research Involving Human Participants

This section may, or may not, be applicable to your project.

This section is not intended to be a duplication of the information you will need to include in ethical approval forms. It is intended to enable you to note the **research data management issues** associated with the ethical considerations of the project. Some examples are given in the table below. These examples are not comprehensive, nor are they applicable to all contexts.

Ethical issue	RDM Actions
<b><i>Informed consent</i></b> <ul style="list-style-type: none"> <li>Need for consent forms plus project information sheet</li> <li>Consent forms must be: (i) checked to ensure they have been completed by all the participants and to note what the individual participants have agreed to (a person might agree to some things on the form, but not all); (ii) stored securely, whether paper and/or electronic; (iii) have a retention period allocated (i.e. how long they will need to be kept)</li> </ul>	Write project information sheet and store Design consent form template and store Check and store signed consent forms. To check the consent forms you could create a record e.g. in a spreadsheet: column for participant code; column for tick that you have their consent form; column for each item on the consent form with a tick/cross/'Y/N' to note the items they agreed/did not agree to.
<b><i>Duty of confidentiality</i></b> This refers to the data given by the participant. <ul style="list-style-type: none"> <li>Data cannot be linked to a named individual unless the participant agrees otherwise</li> <li>Anonymisation procedures</li> <li>Connecting different versions of the data, e.g. after anonymisation</li> </ul>	Look at record of consent forms so you know what individual participants have agreed to. Set up participant codes and key and store Set up the anonymisation protocol and store Anonymise transcripts etc. Obtain the participant's agreement to the anonymised version, if applicable. Store this agreement Destroy audio-records/un-anonymised transcripts (if this was the undertaking given in the consent form/project information sheet). Otherwise, securely store this data. Maintain an audit trail of the different versions of data by e.g. recording the changes, naming conventions etc.
<b><i>Duty of maintaining privacy</i></b> This refers specifically to the personal details of participants such as name and contact details, correspondence. Some personal details do reside in research data, so there is some overlap with duty of confidentiality.	Covered in Section 2.3
<b><i>Consent for how data will be collected, used, processed, kept, disposed of</i></b> Such consent needs to be explicit.	Cover these aspects in the project information sheet Cover in the consent form

<b>Consent for data publication/reuse/sharing</b> Such consent needs to be explicit.	Cover these aspects in the project information sheet Cover in the consent form
Other (please specify)	

If a formal ethical approval process is required, give the location of the relevant documents, e.g. ethical approval form, ethical committee correspondence.

### **Further guidance**

UK Data Archive, Create and Manage Data. Consent and ethics <http://www.data-archive.ac.uk/create-manage/consent-ethics>

## **2.3 Data Protection Act (DPA)**

This section may, or may not, be applicable to your project.

The University is subject to Data Protection legislation, which protects the use of personal data, and as such you must remember that your research data may be covered by the terms of the DPA. This personal data could be collected separately, e.g. a list of participants with their contact details, or be collected during the data collection phase, e.g. a person describing personal details during an interview. If personal data is collected the requirements of the DPA must be observed. These include the explicit consent of participants in accordance with the requirements for collecting and using sensitive personal data. Individuals have rights of access under the Act to their personal information. You may receive a subject access request from a research participant asking what personal information you hold about them and requesting you provide them with a copy of the information. You have a limited time in which to respond and should seek assistance from the University's Records and Information Manager in dealing with any requests. You need to ensure that you have managed your data and records efficiently so you can respond to such a request.

### **Personal data is:**

Data collected either electronically or in paper form

"Personal data means data which relate to a living individual who can be identified –

- (a) from those data, or
- (b) from those data and other information which is in the possession of, or is likely to come into the possession of, the data controller" ICO

### **Sensitive personal data is:**

"Sensitive personal data means personal data consisting of information as to –

- (a) the racial or ethnic origin of the data subject,
  - (b) his political opinions,
  - (c) his religious beliefs or other beliefs of a similar nature,
  - (d) whether he is a member of a trade union (within the meaning of the Trade Union and Labour Relations (Consolidation) Act 1992),
  - (e) his physical or mental health or condition,
  - (f) his sexual life,
  - (g) the commission or alleged commission by him of any offence, or
  - (h) any proceedings for any offence committed or alleged to have been committed by him, the disposal of such proceedings or the sentence of any court in such proceedings.
- The presumption is that, because information about these matters could be used in a discriminatory way, and is likely to be of a private nature, it needs to be treated with greater care than other personal data." ICO

There are eight key **Data Protection Principles**. Personal data should be:

1. Obtained and processed fairly and lawfully
  - effectively this means following these DP principles
2. Obtained and processed only for specified lawful purposes
  - the researcher should be open about what they are going to use the data for, and these uses should not be unreasonable or for an unlawful purpose
3. Adequate, relevant and not excessive
  - the data should be related to the use(s) the researcher has specified, and be the minimum required
4. Accurate and, where necessary, up to date
5. Kept no longer than necessary
  - data no longer needed should be deleted
6. Processed in accordance with the rights of the data subjects
  - there are a number of these rights, e.g. to see a copy of the data held about them, and to have incorrect data corrected/deleted
7. Kept securely and safely
8. Only transferred outside EEA (European Economic Area) with specific protections
  - not all countries have the same level of data protection

There are some exemptions for research data, under controlled circumstances - not a blanket exemption. See the section on 'Use of personal data in research' in the JISC guidance listed below.

Some examples of the RDM issues associated with personal data are given in the table below.

Issue	RDM Actions
<b><i>Individual's consent to process their personal data</i></b>	Cover this in the consent form and project information sheet. See informed consent and other consents in Section 2.2
<b><i>Anonymised data</i></b> This is outside the scope of the DPA.	Decide if you need to collect personal details in the first place. Decide if after anonymisation, and obtaining the participant's agreement to the anonymised version (if applicable), you need to keep personal details. See 'Duty of confidentiality' in Section 2.2.
<b><i>Secure storage of personal data</i></b>	See Section 5.2
<b><i>Secure destruction of personal data</i></b>	See Section 5.2

### **Further guidance**

Northumbria University, Data Protection

<http://www.northumbria.ac.uk/vc/leservteam/ndp/>

Northumbria University, Data Protection and Secure Storage of Research Records

<http://www.northumbria.ac.uk/static/5007/uso/dpgrd.pdf>

Digital Curation Centre, Data Protection <http://www.dcc.ac.uk/resources/briefing-papers/introduction-curation/data-protection>

Information Commissioner's Office, Guide to data protection – definitions, principles and practical examples [http://www.ico.gov.uk/for\\_organisations/data\\_protection/the\\_guide.aspx](http://www.ico.gov.uk/for_organisations/data_protection/the_guide.aspx)

JISC, Data protection code of practice for the HE and FE sectors

[http://www.jisc.ac.uk/publications/generalpublications/2001/pub\\_dpacop\\_0101.aspx](http://www.jisc.ac.uk/publications/generalpublications/2001/pub_dpacop_0101.aspx)

## **2.4 Freedom of Information Act (FoIA)**

This section is applicable to all projects.

The University, as a public body, is subject to Freedom of Information legislation and therefore your research data may be covered by the terms of the Act, which would mean that individuals could request information from your research. This does not mean that information would be made available as exemptions do exist. Where you are re-analysing somebody else's data, you are still responsible for dealing with the FoI request sent to you.

If you get an FoI request you would be assisted by the University's Records and Information Manager. All Freedom of Information requests should be passed on to the Records and Information Manager who will co-ordinate the response with the relevant department or School.

PhD research is a grey area. Any data held by the University (e.g. by supervisors or in packages such as Endnote) would potentially be covered. Any data held by you would also be covered if you are funded by the university or a public body covered by the terms of the Act. It is unlikely that data would be released prior to the completion of the PhD because of the impact this would have on you. To release such data would undermine the PhD structure and would not be in the public interest. However, each case would need to be judged on its own individual merits.

### **How to plan**

- Efficient research data management means that responding to an FoI request can be expedited
- Placing research data/summary findings in the public domain, as applicable/feasible, can pre-empt any such request

### **Further Guidance**

Northumbria University, Freedom of Information

<http://www.northumbria.ac.uk/vc/leservteam/foi/>

JISC, Freedom of Information and research data: Questions and answers

<http://www.jisc.ac.uk/publications/programmerelated/2010/foiresearchdata.aspx>

## 2.5 Other Legal Issues

This section may, or may not, be applicable to your project.

There may be other legal issues related to your research. Some examples are given in the table below. These examples are not comprehensive, nor are they applicable to all contexts.

Legal Issue	RDM Actions
<b>Official Secrets Act</b> It is an offence for a person who has signed the Official Secrets Act to divulge, without permission, any information/documents related to their official work	Keep data secure and confidential
<b>Animals (Scientific Procedures) Act</b> Use of animals in research is regulated and requires licenses	Store the required licenses, and keep them up to date
<b>Control of Substances Hazardous to Health Regulations</b> Employers are required by law to control substances that are hazardous to health.	For the employer: Store the risk assessment document Store the safety data sheets and keep them up to date Make the safety data sheets accessible to researchers, technicians etc. Store any required permits and keep them up to date Keep records of the results of monitoring procedures and health surveillances Store the emergency plans and keep them up to date Make the emergency plans accessible to researchers, technicians etc.
<b>Environmental Information Regulations</b> Public authorities (which include universities) are required by law to provide information about the environment which could be affected by their activities, in a similar fashion to FoI requirements.	If you receive an Environmental Information request contact the University's Records and Information Manager immediately

### Further Guidance

Health and Safety Executive, Control of Substances Hazardous to Health (COSHH), <http://www.hse.gov.uk/coshh/>

Home Office, Research and testing using animals, <http://www.homeoffice.gov.uk/science-research/animal-research/>

Information Commissioner's Office, Environmental Information Regulations, [http://www.ico.gov.uk/for\\_organisations/environmental\\_information.aspx](http://www.ico.gov.uk/for_organisations/environmental_information.aspx)

## 2.6 Intellectual Property Rights (IPR)

This section is applicable to all projects.

Intellectual property rights include:

- Patents
- Trademarks
- Designs
- Copyright
- Confidentiality agreements

Who owns the intellectual property right (IPR) to the data? It could be a number of different people/organisations, e.g.

- yourself
- other researchers in the team
- the University
- a funding body
- participants - they hold the copyright to their own words
- an organisation that provided services to the research activity, e.g. equipment, study organisation site
- another person etc.

If more than one person owns the IPR, what agreement do you have on how this is to be handled?

Get the IPR issue set out within the terms of the project agreement, either fully or in outline, and if necessary, have the IPA further discussed and agreed by the parties. For multi-partner projects, this could be covered in a consortium agreement. It is important to achieve an appropriate balance between concern for legal implications and getting the research done.

For participants, include a statement in the consent form asking them to pass their copyright to the research team.

### **Further Guidance**

Northumbria University, Financial Regulations, v18, 20 May 2012

<http://www.northumbria.ac.uk/static/5007/finpdf/financialregs.pdf>

*27.14. The intellectual property rights for any idea, design or invention that arises from work undertaken by a member of staff whilst employed by the University, shall as provided by his or her contract of employment, normally belong to the University. Further information and guidance is available from Research and Business Services.*

Northumbria University, Regulations for the degrees of Master of Philosophy (MPhil) & Doctor of Philosophy (PhD) from 1 September 2013,

[http://www.northumbria.ac.uk/static/5007/graduateschool/regs\\_mphilphd13.pdf](http://www.northumbria.ac.uk/static/5007/graduateschool/regs_mphilphd13.pdf)

*11.2. The copyright of the submission rests with the student.*

*11.3. As a condition of enrolment, the student agrees that the University shall own any Intellectual Property (IP) that may result from his/her research activity ... In the case where a student is funded wholly or partly by a third party external to the University, the supervisor(s) and the student should clarify intellectual property ownership issues with the third party at the outset of the research project.*

UK Intellectual Property Office <http://www.ipo.gov.uk/>

## 2.5 Other Agreements

This section may, or may not, be applicable to your project.

Stakeholders in a research project, e.g. different researchers, funding bodies, partner organisations etc, will have expectations about the way the research will be conducted. Establish agreements surrounding these expectations.

Expectations might cover such topics as:

- funder/partner/sponsor organisation might want to have a copy of the anonymised data
- provider of specialised equipment or services might want a non-disclosure agreement to protect their commercial information
- data reuse (see Section 3.1)
- data sharing (see Section 3.2)
- publication of outputs (see Section 6)

The project might obtain services from other people/organisations. In that case the project may want/need them to meet certain requirements, e.g.

- transcribers/translators of participant data: they would need to sign an agreement that they would keep the data secure and confidential
- providers of IT services, such as databases, cloud computing: you would need to ensure that they do not claim any IPR over the data

## 3. Data Reuse and Sharing

You will come across different usages of the terms 'data reuse' and 'data sharing' so it is useful to provide some definitions here.

- Data use is using the data for the current research project.
- Data reuse is using data from one research project for another purpose, or another research project.
- Data sharing is making data available to other researchers for them to reuse it.

### 3.1 Data Reuse

This section may, or may not, be applicable to your project.

For the purposes of this DMP we have defined data reuse as the use of the data from the current project by yourself and/or the other researchers in the project team for a **different** project. If you are going to reuse data then you need to do 3 main things:

- obtain all the necessary permissions to do so, and establish any restrictions on how the data can be reused
- obtain all the necessary contextual information to make the data meaningful
  - even though you will be reusing the data you cannot rely on your memory to provide the necessary context
- store the data, preserve it and ensure it is accessible (see Section 7.2 for details)
  - the methods for doing this will be far less onerous than for data sharing in Section 3.2, e.g. you could keep the data on your personal drive or a shared drive within the University's IT system

## 3.2. Data Sharing

This section may, or may not, be applicable to your project.

For the purposes of this DMP we have defined data sharing as making your data from the current project available to other people.

There could be requirements for you to share data, e.g.:

- funding bodies, particularly the Research Councils, might require that data is shared, unless there are good reasons for not doing so

There could be other reasons why you might want to share data, e.g.:

- the data is scientifically, socially, or culturally significant
- the data is the only or most complete source of the information it contains
- it would not be feasible, or financially viable, to replicate the data

If you don't want to share your data it is useful to note the reasons why, e.g. you might need to justify this decision to a funding body. Some reasons for NOT sharing data include:

- you might be able to share some parts of the data not others
- you might be able to share anonymised/summarised/synthesised data, but not raw or fuller versions of the data
- ethical reasons, e.g. the data is too sensitive or confidential; you have not obtained permission for sharing from the participants
- you have a non-disclosure agreement
- quality-related issues
- epistemological/methodological issues

If you are going to share data then you need to do 3 main things:

- obtain all the necessary permissions to do so, and establish any restrictions on how the data can be reused
- obtain all the necessary contextual information to make the data meaningful
- store the data, preserve it and make it accessible (see Section 7.2 for details)

### Further guidance

Research Councils UK, Common principles on data policy,  
<http://www.rcuk.ac.uk/research/datapolicy/>

*Publicly funded research data are a public good, produced in the public interest, which should be made openly available with as few restrictions as possible in a timely and responsible manner that does not harm intellectual property.*

## 4. Data Collection and Analysis

### 4.1 Using Existing Data

This section may, or may not, be applicable to your project.

National and international bodies are emphasising the importance of data sharing, which leads naturally to this data being reused in new projects. Some research funders are now asking applicants to justify why they are seeking funding to collect new data rather than re-analysing existing data.

If you are using existing data this has two main implications:

- the data is managed by the source; your responsibility is to manage the new data/outputs you derive from it
- there may be conditions about how you can use the data

## 4.2 Creating and Capturing New Data

This section may, or may not, be applicable to your project.

Think how you will create and capture your new data, what data processing is required, what equipment/hardware/software you will be using and the associated file types.

Examples of data creation/capture methods and associated data processing are given in the table below. These examples are not comprehensive, nor are they applicable to all contexts.

Data Creation Methods	Data Capture Methods	Data Processing
interviews/focus groups/observations	writing notes on paper writing notes on an electronic device making audio/visual record	transcription anonymisation transfer of data into software package for analysis, e.g. NVivo
questionnaires	writing answers on a paper questionnaire filing in an electronic questionnaire	typing data into a software package for analysis, e.g. a spreadsheet/SPSS transferring data into a software package for analysis, e.g. a spreadsheet/SPSS
experiments	recording analogue/digital data via a specialised device	data cleansing transferring data into database for analysis

As well as the data itself, you will be creating associated materials which will also need to be managed. Examples of associated materials include:

- project information sheets
- consent forms: template and signed copies from participants
- correspondence with participants and other people associated with the project
- researcher diary/log/lab notebook
- protocols, e.g. overall research methodology; anonymisation protocol
- research instruments, e.g. interview schedule, questionnaire, experimental set up

Knowing what data and associated materials you will be creating will enable you to make decisions about, e.g.:

- the location where data capture takes place. If it's in the field you will need portable devices and the issue of data security arises. See Section 5.2
- the different types of data you will be creating and the different requirements they have for their management
- the kinds of protocols/standard operating procedures you will need to develop, e.g. transcription protocol, anonymisation protocol
- how to maintain an audit trail between the different versions of the data as it undergoes various processes
- version control for the different stages of processing a data set

- the folder structure for organising your files

You will be using a range of software and hardware, and generating different file types. Keeping brief details of these will enable you to check things like:

- do you have access to all the hardware/software you need
- can everyone involved in the project use these file types or would files need to be formatted differently, e.g. a partner organisation might not have access to a particular piece of software, or only have an earlier version
- keeping up with changes in software that might make your old files unreadable

You will be producing files of many different formats, e.g. electronic text/word document (.doc/.docx); MS Excel (.xls/.xlsx) or MS Access (.mdb/.accdb) for capturing questionnaire responses; NVivo (.nvp/.rtf/.txt) or ATLAS.ti (.hpr6) for analysis, etc. Think about the reasons why you are using these formats, e.g. they are in widespread use, the University has the relevant software licences or they're accepted standards in your field, you have staff expertise in their use, etc. However use of proprietary software and certain types of formats can restrict your ability to reuse and share data in the future. So consider if you can use, or export to, open formats, accepted standards, etc.

### **Further guidance**

Northumbria University, Faculty of Engineering and Environment, DATUM in Action - Supporting researchers to plan and manage their research data. Folders and files – Guidance.

[http://www.northumbria.ac.uk/sd/academic/ee/work/research/clis/information\\_records\\_management/rmarea/datum/action/outputs/](http://www.northumbria.ac.uk/sd/academic/ee/work/research/clis/information_records_management/rmarea/datum/action/outputs/)

- This guidance covers: folder structure, file naming, version control

## **4.3 Data Analysis**

This section is applicable to all projects.

Think how you will analyse your data, what equipment/hardware/software you will be using and the associated file types.

The guidance given under Section 4.2 above is also applicable here.

## **4.4 Contextual Information**

This section is applicable to all projects.

If data is to be usable by yourself during the project and afterwards, and if it is to be shared, then it needs to be meaningful and easily understood and interpreted. Therefore the data needs to be described and annotated, and contextual information and documentation need to be produced. This documentation/description is needed at many levels:

- the project level: e.g. project proposal; methodology, research context, categorisation of participants, data collection/analysis protocols, analysis coding schemes; data collection tools (e.g. blank questionnaire, interview schedule, equipment and experimental set up); findings (see Section 6.)
- data level: e.g. database structure; names, labels, codes, descriptions of variables/data items. Contextual information at this level will overlap with metadata (see Section 4.5)

It is best to create such documentation from the outset of the project: relying on your memory to reproduce such information at the end of a project is risky and inadequate.

### **Further guidance**

Northumbria University, Faculty of Engineering and Environment, DATUM in Action - Supporting researchers to plan and manage their research data. Metadata guidance. [http://www.northumbria.ac.uk/sd/academic/ee/work/research/clis/information\\_records\\_management/rmarea/datum/action/outputs/](http://www.northumbria.ac.uk/sd/academic/ee/work/research/clis/information_records_management/rmarea/datum/action/outputs/)

- This guidance covers: contextual information, metadata

## **4.5 Metadata**

This section is applicable to all projects.

Metadata is structured information/data that makes your research data (and associated materials and documents) meaningful and usable. It covers both electronic and physical items, and is intended to be read by humans and/or computers. It includes:

- descriptive metadata i.e. information used to search for, identify and locate the data, e.g. author, title, date, keywords
- structural/technical metadata i.e. information that enables the objects to be used and manipulated, e.g. hardware/software/equipment, file formats, database elements
- administrative metadata i.e. information that enables the data to be managed, e.g. versions, ownership/IPA, licensing
- use metadata i.e. information to manage access to the data, e.g. access rights, usage tracking
- preservation metadata i.e. information for managing preservation activities, e.g. actions such as migrations

For conducting a research project descriptive, structural/technical and some administrative metadata (particularly versions) is usually all that is required. Administrative, use and preservation metadata are required for long term storage and for data sharing and if data is placed into a repository such activities are undertaken by the repository.

Some metadata is obtained automatically, e.g. by the data collecting equipment used (such as a digital camera containing a clock, calendar and GPS which can therefore generate the date, time and location of an image and embed this information within the image; scientific instruments can provide very detailed, sophisticated metadata), through the design of the database holding the data (such as the attributes of fields, rows, columns etc.), from inheriting document/folder properties (if you have access to systems in your organisation that can set these up). Other metadata must be created manually - this is particularly true for file-based systems - either within the file itself or in associated documents/systems.

Find out what systems/mechanisms/equipment you have access to that can collect/create automatic metadata. For manually created metadata decide what you really need - manually creating metadata is a burden, and it is pointless to require metadata that the researchers will not produce on a sustained basis throughout the project. If you decide to pass your data on to a repository for long term storage/sharing then you will be required to provide specified metadata. It may be helpful to use metadata standards. Such standards can be generic, e.g. the Dublin Core for describing resources, or subject specific, e.g. Directory Interchange Format for earth science data.

## **Further guidance**

Northumbria University, School of Computing, Engineering & Information Sciences, DATUM in Action - Supporting researchers to plan and manage their research data. Metadata guidance.

[http://www.northumbria.ac.uk/sd/academic/ee/work/research/clis/information\\_records\\_management/rmarea/datum/action/outputs/](http://www.northumbria.ac.uk/sd/academic/ee/work/research/clis/information_records_management/rmarea/datum/action/outputs/)

- This guidance covers: contextual information, metadata

Directory Interchange Format <http://gcdm.gsfc.nasa.gov/add/difguide/index.html>

The Dublin Core Metadata Initiative <http://dublincore.org/> and <http://dublincore.org/documents/dces/>

## **5 Short-term Data Storage**

Short term data storage covers the duration of your project. See also Section 6, Long term data storage.

### **5.1 Data Storage**

This section is applicable to all projects.

It is likely you will need both electronic and physical storage, so you have to ensure that you have sufficient facilities/equipment and space for the anticipated amount of items you will generate during the project.

Though you will be using systems within your university, you will also be using portable devices during data collection activities/field trips, and may also have items on computers at home or in systems in other organisations. Think about all the locations where your data will be stored. A lot of the decisions and actions about data storage during a project is the responsibility of the researchers. Ensure that all members of the project team understand their responsibilities for data storage, security (Section 5.2), retention and destruction (Section 5.3), and preservation (Section 5.4).

## **Further Guidance**

UK Data Archive, Create and Manage Data. Storing your data <http://www.data-archive.ac.uk/create-manage/storage>

- This guidance covers: backing-up data, storing data, data security, transmitting & encrypting data, file sharing, data disposal

### **5.2 Data Security**

This section is applicable to all projects.

You need to consider the security of your data from many different perspectives:

- all the activities you are undertaking: different approaches may be needed for different activities
  - during data collection
  - during data transfer
  - whilst stored
  - during data access
  - at data destruction

- the nature of the data: different levels of risk apply to different categories of data
  - personal/sensitive data needs a far higher level of security

The security of data on the University system is well managed by the organisation: access is controlled by id/passwords; the data is regularly backed up; the data is protected by firewalls and virus checkers. The researcher however has responsibility for ensuring the security of data on home computers, e.g. by using a virus checker and regularly backing up the data onto another device.

### **Further Guidance**

Northumbria University, Faculty of Engineering and Environment, DATUM in Action - Supporting researchers to plan and manage their research data. Information security guidance.

[http://www.northumbria.ac.uk/sd/academic/ee/work/research/clis/information\\_records\\_management/rmarea/datum/action/outputs/](http://www.northumbria.ac.uk/sd/academic/ee/work/research/clis/information_records_management/rmarea/datum/action/outputs/)

UK Data Archive, Create and Manage Data. Storing your data <http://www.data-archive.ac.uk/create-manage/storage>

- This guidance covers: backing-up data, storing data, data security, transmitting & encrypting data, file sharing, data disposal

## **5.3 Data Retention and Destruction**

This section is applicable to all projects.

During the project you may want to dispose of some data/materials, e.g.

- personal data that you no longer need - to meet the requirements of the Data Protection Act
- raw data in audio files and original transcripts - to protect participant confidentiality
- transitory documents - disposing of these clears clutter on your drive, in your email inbox, on your desk

To decide what you need to keep and what you can discard requires appraisal (see Section 7.1). As part of this appraisal process, determine what items can be destroyed during the project. Ensure you have procedures/equipment to destroy confidential data securely.

### **Further Guidance**

JISC Infonet, HEI Records Management. Guidance on Managing Transitory Records

<http://www.jiscinfonet.ac.uk/partnerships/records-retention-he/managing-transitory-records>

UK Data Archive, Create and Manage Data. Storing your data <http://www.data-archive.ac.uk/create-manage/storage>

- This guidance covers: backing-up data, storing data, data security, transmitting & encrypting data, file sharing, data disposal

## 5.4 Data Preservation

This section may, or may not, be applicable to your project.

Electronic data is subject to continuous change: hardware/software become obsolete on a regular basis; digital media break down in a few years. Therefore electronic files need to be preserved in some way so they are still readable/usable when required. It is likely that over the period of a PhD project or a research project funded for a number of years that there will be some changes to your IT that might affect your electronic data. Therefore, preservation is best considered from the outset of a project, not left too late when there may be no data able to be preserved. Such preservation decisions are also needed if data is to be retained for a long period after the end of the project, or shared. The need for preservation actions during a project will depend on its length: a project of 6 months duration is unlikely to need to take actions; a project of a few years certainly will. A preservation plan would include:

- making someone responsible for this task
- watching out for technological changes that might affect your data
- using data formats that are better for preservation, e.g. non-proprietary formats or formats based on open standards
- migration of data if you need to, i.e. copying your data into formats supported by the new software/hardware; copying your data onto new media.

### **Further Guidance**

Digital Preservation Coalition, Publications <http://www.dpconline.org/publications>

## 6. Dissemination

This section is applicable to all projects.

This section covers the outputs and publications from your project, i.e. the findings/results. The publication of the data itself is covered in Section 3. There may be requirements on what you can publish from your project, and how it should be published, e.g.:

- some reports can be so commercially sensitive that they cannot be placed into the public domain
- some funders/governments might require open publication of research results

Document any requirements about the publication of your results, and any agreements on publications you reach with your research colleagues.

### **Further guidance**

Creative Commons UK <http://www.creativecommons.org.uk/>

Research Councils UK, Policy on Access to Research Outputs  
<http://www.rcuk.ac.uk/research/outputs/>

## 7. Long -term Data Storage

Long term data storage covers after the end of your project. See also Section 5, Short term data storage.

## 7.1 Data Retention and Destruction

This section may, or may not, be applicable to your project.

To decide what you need to keep, and for how long (retention period), and what you can discard requires appraisal. Think about:

- the laws and regulations which affect your data, i.e. what you are allowed, or not allowed, to keep
- what your university/sponsors/funders expect you to keep, for how long and where
- the minimum you need to keep for your findings and publications to be supported over time
- what you want to do with the data – now and in the future after the end of the project
- what other people might want to do with your data - and what you would want them to be able to do

Look at the decisions you made in Section 3.1 (Data reuse) and 3.2 (Data sharing).

There are general criteria which could help you make such decisions:

### 1. Relevance to Mission

- Does the item content fit the organisation's remit and priorities, including any legal requirement to retain the item beyond its immediate use?

### 2. Scientific, Social, Cultural, Historical Value

- Is the item scientifically, socially, or culturally significant? You need to anticipate future use, from evidence of current value.

### 3. Uniqueness

- To what extent is the item the only or most complete source of the information it contains?

### 4. Potential for Redistribution

- Is the item authentic (i.e. what it says it is), with integrity (unchanged), and usable? Does it meet IPR and ethical requirements?

### 5. Non-Replicability

- Would it be feasible, or financially viable, to replicate the item?

### 6. Economic Case

- What is the cost of managing and preserving the item, and does the value of the item justify this cost?

### 7. Full Documentation

- Is the metadata and contextual information needed to find, access and reuse the item comprehensive and correct?

Document the decisions you make, and monitor that you have acted on them.

### **Further guidance**

Northumbria University, Records Retention Schedule, Research  
<http://www.northumbria.ac.uk/vc/leservteam/recman/RRS/>

## 7.2 Data Storage

This section may, or may not, be applicable to your project.

If you are going to retain your data long term then you must decide where it will be kept.

If you are going to place it in a repository, e.g. the University repository or a national repository such as the UK Data Archive, then your responsibility ends once they have taken over its management. However, you will have to prepare the data for deposit, e.g. ensuring you have all the necessary consents and permissions; ensuring you have sufficient contextual information and metadata; ensuring you meet the repository's requirements, e.g. formatting, etc.

If you are going to keep responsibility for your data long term, e.g. keeping it on your computer drive in the University, or at home, or on a Website, then you must put into place mechanisms to ensure its security, control access to it, and preserve it.

### **Further Guidance**

UK Data Archive <http://www.data-archive.ac.uk/home>

UK Data Archive, Create and Manage Data. Storing your data <http://www.data-archive.ac.uk/create-manage/storage>

- This guidance covers: backing-up data, storing data, data security, transmitting & encrypting data, file sharing, data disposal

## **8. Resources**

This section may, or may not, be applicable to your project.

Think about the resources you need to be able to implement your DMP, e.g. software tools, equipment, computing storage space, storage media. Plan for how you will obtain these resources.

---

<sup>i</sup> Digital Curation Centre (DCC) DMP Template <http://dmponline.dcc.ac.uk/>  
and Help <https://dmponline.dcc.ac.uk/help>