DATA MINDFULNESS:

MAKING THE MOST OF YOUR DISSERTATION

RESEARCH DATA SERVICE

data-support@ed.ac.uk

Candela Sanchez

csanchez@exseed.ed.ac.uk

THE UNIVERSITY *of* EDINBURGH

is information SERVICES

THE UNIVERSITY *of* EDINBURGH

is information SERVICES

THE UNIVERSITY
of EDINBURGH

information SERVICES

PREFACE:

Hi there!

This workshop and booklet will, hopefully, help you start thinking about your dissertation journey. My name is Candela and I currently work for the Research Data Service putting together these training materials, although I am doing a PhD as well. I remember when I did my UG dissertation and how lost I was; and now I hope this resource will make you feel more prepared knowing some tips and information about what lies ahead, and also what things you should be aware of.

Both the workshop and this booklet are structured following the journey you will go through, from thinking of a research question to writing up and dealing with your dissertation after submission. Keep in mind that this resource has been designed to suit all students from the University, and so there may be sections that are more or less relevant to your specific discipline. Additionally, this is only a starting point to get you thinking about your dissertation. The Library and Information Services have more specific and in-depth courses that you may want to have a look at.

Best of luck with your dissertation! And have fun!

1. WHAT IS A DISSERTATION?

Yes, what is a dissertation? And how is it different from a traditional essay you have written for one of your courses? Understanding this, even though it may seem unimportant, is *crucial* for conducting and writing an excellent dissertation. So, let us unpick this by parts:

First, what is a dissertation? A dissertation is about *doing* research, rather than *writing about* research that others have conducted in order to answer a question and construct an argument; and, although they may take many different forms depending on your degree and your topic, most UG dissertations do share that:

1. The student chooses the topic and also approach to the topic of their work (unlike a course essay).

2. This piece of work will be carried out in an individual basis — of course, you will have guidance from your supervisor.

3. As I have mentioned already, a dissertation is about doing research, and so a UG dissertation will have a substantial component of research — whether this means collecting primary data, analysing existing data, or critically analysing text materials.

4. And finally, it is long! Much longer than a traditional course essay. Most UG dissertations are around 10,000 words. And this means that it will also be much more in depth too.

1.1. WHAT DOES A DISSERTATION LOOK LIKE?

Your dissertation will be structured differently depending on your discipline and your topic:

THEORY OR TEXT BASED DISSERTATIONS	ARTS, HUMANITIES AND SOCIAL SCIENCES	SOCIAL SCIENCES (TRADITIONAL) AND SCIENCES
1. Introduction	1. Introduction	1. Introduction
	2. Literature Review	2. Literature Review
	3. Research Design and Methodology	3. Research Design and Methodology
2. Main Body		4. Findings
	4. Main Body	5. Discussion
3. Conclusion	5. Conclusion	6. Conclusion

Dissertations are structured in specific ways, depending on your discipline and on your topic; and this structure will be more or less prescribed (see the table above).

The structure of your dissertation reflects the type of research that you are conducting. For example, in the sciences and more traditional social sciences, dissertations tend to have a more prescribed structure; this includes a chapter that describes your research design and methodology; a chapter that describes your findings, without any interpretation from you; and a chapter or set of chapters in which you discuss and interpret those findings (see the third column in the table above). Now, this way of structuring your dissertation is built on the assumption — as most science and some social science research is — that you, as the researcher, are able to collect objective data and are able to describe those free from any values, judgements, social preconceptions and theoretical interpretations.

On the other hand, other disciplines and other research projects may have different assumptions about the research process and the way in which research data is collected or produced; and this is reflected in the way those dissertations or thesis are structured. For example, many social science or humanities projects would sit on the assumption that describing truly objective data is unattainable; and therefore, those projects are likely to have a set of chapters that both describe and interpret their findings (see the second column in the table above).

Theoretical dissertations have the least prescribed structure, as you will not normally have a methodology or results to write about. Your dissertation will be structured in chapters that all discuss literature and build towards your central argument. For example, my MSc dissertation was around the concept of power. It was a political theory dissertation in which I discussed the concept of 'empowerment' within the context of power theory. So, my dissertation structure took the reader from an overview of the literature on 'empowerment', to the theoretical problems with this concept, to a set of chapters in which I developed a theoretical solution to the 'problems' I had stated earlier.

Now, the key to a dissertation, no matter which type you are writing, is that you will have a) a research question; b) you will develop an argument to answer that research question; and c) you will base that argument on evidence that you will show and discuss. That evidence - or data - encompasses a very wide range of items and this is what we will look at next.

2. WHAT ARE 'DATA'?

Research data encompasses a much broader range of items than we normally would think. Every single student who takes on a research project will produce and use research data, from those doing pure theoretical work, to those designing and running lab experiments, to those conducting art-based research, to those doing quantitative or qualitative research. All of us, no matter the project will produce some form of research data. And this workshop — and this workbook — aims at taking you along the research journey and showing you what issues you should be mindful of when dealing with your own research data.

So, before we go any further into this course, here are some examples of data that you may be dealing with at each stage of your research project:

THE RESEARCH JOURNEY	EXAMPLES OF DATA
Finding your Research Question	Literature Notes that you make Literature reviews Systematic reviews
Research design	Research plan documents Methodology notes Research diaries Consent forms Ethics forms Recruitment forms/posters/ads Experiment designs
Data collection	Observations Archival material Texts Still images, video, and audio Art artefacts Simulation data, models, software Sketches, diaries, notes Temporal data, biological data, specimens, land features Protein structures, sensor readings Historical records, census records, opinion polls Economic indicators

THE UNIVERSITY *of* EDINBURGH

is information SERVICES

THE RESEARCH JOURNEY	EXAMPLES OF DATA
Data production/handling/analysis	Transcripts Databases Notes Research diaries Texts Statistical results Theory notes
Write-up	Sections of writing Different versions

These are only some examples, depending on the kind of project that you will be doing these will change, and it may also be that some aren't covered here. It may also be that, for example, if you are conducting a theory based dissertation you will mainly be dealing with texts and not thinking about data collection at all. Yet, you will still produce data — in the form of text, notes, perhaps mind maps — and knowing how best to deal with those is also part of this workshop.

3. DEALING WITH LITERATURE

The first step in conducting your dissertation research is likely to be conducting a literature search. This will be both part of the research process — to enable you to find a gap in the literature, and thus justify your research; but also it will likely become either a chapter in your dissertation, or if you are conducting a theory dissertation, this will be most of the content of your dissertation.

When dealing with literature, it is too easy to download an article, read it, take notes, and then misplace it — or even forget which article those notes are about. I know I have definitely done that, and then cursed myself for not having a more thoughtful approach to dealing with literature. Another issue that seems to come up time and time again with students that I have taught is referencing; and by this I mean a) having a system to store and organise your sources, whatever those are; and b) referencing those correctly.

To avoid these issues:

1. Think methodically about your literature search and storage:

Create a particular folder or folders for your literature and/or other sources. Now, this may be hard, because you may not know in advance which research areas you will be getting into, and the structure and organisation of your literature is likely to change as you read more and get more of a sense of, say, how all of your articles should be grouped together. However, thinking about this in advance will be beneficial in two ways: a) it will mean that you can keep track of the literature you have read or thought may be relevant for your research project (even if you don't actually read it straight away, save it!); and b) thinking about how to organise your literature will actually help you in writing your literature review. This is because the way in which you organise your literature — and I mean your files, in folders — is very much part of the process of making sense of the links and relations between those very papers; and that is at the core of writing a good literature review. Think of it like making a mind map of everything you are reading.

2. Create a strategy for file naming - and stick to it!

When you download an article from an online journal, my advice is to always save it in the folder you designated for literature, and using always the same name structure. One I recommend my students is:

Author_Year_Title of the article_KEY WORD_KEY WORD_KEY WORD

The reason for this is that creating a naming strategy will mean that you can a) find articles quickly, without needing to open each file to see what it is all about; and b) adding the KEY WORDS at the end of the file name

THE UNIVERSITY
of EDINBURGH

is information SERVICES

will also enable to use your search function and sieve through all of your literature fast when you are looking for specific themes.

3. Start using referencing managers early on!

Referencing managers are pieces of software that allow you to collect, store and manage your references in searchable folders; insert in-text citations (as many of them will plug into Microsoft Word); and generate automatic bibliographies too.

They also allow you to:

- Capture related PDF's, web pages, files, or images or link to available full-text.

- Add personal notes and index PDF full-text.

- Format citations in a wide range of reference styles e.g. Harvard, Vancouver.

- Create user groups and share references for class and other collaborative research work.

- Take notes on articles and save them in your collection of citations.

If you are using a University-supported computer, it will have EndNote installed and so you can use it for free. If you need a reference manager for your own computer, you can:

1. Purchase a discounted version of EndNote here:

    https://www.ed.ac.uk/information-services/computing/desktop-personal/software/main-software-deals/endnote/erm-personalcopies

2. You can use EndNote's free web version. Although this only allows you to store and manage references; not the citing while you write option.

3. You can use an open source referencing manager. There are some very open source – and free – options such as Mendeley or Zotero, which you can download on your personal laptop or PC.

If you have never used a reference manager and/or would like to learn more about them, the library runs courses specifically on using EndNote; you can find these courses and sign up on Learn. The University also has an in-depth workbook you can have a look at: https://www.ed.ac.uk/information-services/help-consultancy/is-skills/catalogue/text-catalogue/endnote-managing.

4. RESEARCH DESIGN

By this point of the research journey you will have a) come up with your Research Question; and b) will have become familiar with the research area that you would like to contribute to. And yes, I know, your question may be only a tentative one and will probably change during the course of your dissertation work. Despite that, it is important to have a question, even if you are aware that it is not final, because that will allow you to move onto designing your research project. And that is the next step in your research journey.

Research design and methodology are extremely broad terms, and although we may associate them with scientific or social science dissertations; they do apply to all research projects. I like to think of research design as, simply, the thinking that goes behind deciding what you actually need in order to answer your research question. And this is actually a really important decision and will vary depending on your research theoretical stance. For example, my PhD thesis asks (to put it simply) what is the relationship between the experience of panic attacks and the places where people experience them. Now, I chose to answer this question using a small amount of in-depth qualitative interviews; however, someone else may have chosen to answer this question using questionnaires and survey data; or instead, looking at the design elements of those places; or looking at it from a more clinical or medical perspective. And this applies also to theory based dissertations. You will need to make choices as to what theories are relevant to construct your particular argument.

All of these choices are what determine your research design, and therefore what data you will use in your dissertation project. And this, in turn will determine where to source that data from — whether from texts, data repositories, existing records, or produce it yourself. Have a look at the table[1] below; it gives you a sense of different types of data and also their possible sources.

Sciences	- Meteorological data, ground water data, sensor readings, historical records - X-rays, clinical case studies - Chemical structural data, crystal structure, molecular calculations - Spectral surveys - Specimens, biodiversity surveys - Experiment data, observations, calculations	Generate your own data Obtain from other researchers Data repositories

[1] Table adapted from Borgman, C. (2014) 'What are Data?' [PowerPoint presentation]. Available at: http://www.slideshare.net/ntunmg/what-are-data (Accessed: 22 March 2019)

THE UNIVERSITY *of* EDINBURGH

is information SERVICES

Social Sciences	- Opinion polls - Surveys - Interviews - Mass media, social media - Laboratory experiments - Field experiments - Fieldwork notes - Demographic records - Census records - Voting records - Economic indicators	Generate your own data Obtain it from other researchers Data repositories Existing records
Arts and Humanities	- Newspapers - Photographs, video material - Letters - Diaries - Literature: books, articles - Church records - Court records - Maps - Art artefacts - Historic artefacts	Libraries Archives Museums Public/corporate/government records Data repositories

This is not an exhaustive list, of course, but I hope that it gives you an idea of how extensive data can be, and also, that the places for sourcing data are almost as extensive. In the next section of this handbook I will go through some of the issues that you may want to think about when managing your dissertation data; however, before diving into that, there is two pieces of advice that I would very strongly encourage to take with you:

1. Be **mindful and thoughtful** about the data you choose to answer your research question. Think about why you are choosing such type of data and what are the consequences for the way in which you will design your research methodologies. One of the key aspects of an excellent piece of research — and this applies to dissertation projects to PhD projects and to very large research projects too — is coherence. Make sure that your research is coherent. And what I mean by this is, make your choices about your research design and data consciously and justify them.

2. And that brings me to my second point; which is, **document all of this process**. Document your draft research design. Save copies of experiment designs, or interview plans, even if you don't end up using them. The reason for this is that, when you come to write your methodology or research design chapter

in your dissertation, all of this will help you describe and explain your research process more thoroughly. And this, in turn, will produce a dissertation that is strong and transparent. And these are the foundation of any good piece of research.

5. DATA COLLECTION AND ANALYSIS

The next step will be to 'collect' your data. Whether this is accessing secondary material, generating your own data or using only literature. This section will teach you how to manage your data well; and in the cases where you are generating your own data, I will also discuss some of the issues that you must pay attention to if you are generating sensitive data.

5.1. STORING YOUR DATA

5.1.1. STORAGE: M: DRIVE

In the same way that I advised you to think methodically about how to deal with your literature, it is important that you think carefully about where you will store your data - whether that is simply your own typed notes, paper notes, artefacts, or a complex set of data. All of those will be an important part of your dissertation work and avoiding any possibility of data loss is crucial. So, where do I recommend that you store your dissertation data?

The best option is storing it all in your university network drive (that is your M: student drive). You will be familiar with this by now. As you will know, all students are given up to 2GB of data storage that is managed by your own School and this is mapped as the M: drive. The benefits of using this storage is that these servers are secure (and we will look at why this is important in the next section of this handbook), and periodically backed up; so if you accidentally deleted a file it would be easily recoverable. Documents and files saved on networks storage are also able to be rolled back to a previously saved version. Now, if you normally work from the University's managed computers accessing your M: drive is very straight forward. If, however, you tend to work from home or on your personal laptop there is a way to link your personal computer to your M: drive. This may seem a little cumbersome at first, but you only need to do this once. After this initial set up, you will be able to very quickly access your M: drive from anywhere, as long as you have an internet connection.

Remotely connecting to your M: drive

First of all, you will need to a) set up VPN access in your computer; and b) find your own personal folder address. Once you have those two, connecting to your M: drive is pretty straight forward.

a)   Connecting to your VPN:

Open the link below. Click on your operating system, under 'Built-in VPN clients' and follow the instructions:
https://www.ed.ac.uk/information-services/computing/desktop-personal/vpn/vpn-service-using

b) Finding your own personal M: drive address:

Open the link below:

https://myadinfo.is.ed.ac.uk/MyInfo.aspx

This will redirect you to the EASE log-in page. Enter your UUN and password, and once you have done this, you will be shown your account information, including your Active Directory Paths - for both Windows and for other providers (Mac, Linux etc). Make a note of this as you will need it in the following set of steps.

WINDOWS 10	MAC OS X
1. Connect to the University's VPN	1. Connect to the University's VPN
2. Open any folder	2. Switch to the Finder (e.g. click anywhere on the desktop)
3. On the left column, right click on 'This PC' and then click 'Map Network Drive'	3. Click on 'Go' in the menu bar, then click on 'Connect to Server'
4. Pick an available letter	4. Enter your folder address (this is your Active Directory Paths; it will start with smb://)
5. Enter your folder address (this is your Active Directory Path; it will start with \\)	5. Authenticate with your username in the form of UUN@ed.ac.uk
6. Authenticate with your username in the form of UUN@ed.ac.uk	

Once you have connected to your M: drive, you will be able to see it and access it from your own personal computer or laptop.

5.1.2. STORAGE: CLOUD-BASED OPTIONS

All University students have access to Microsoft OneDrive. This is a cloud-based storage service integrated into Microsoft's Office 365 suite of applications. Full details on accessing and using OneDrive can be found in the link below:

http://www.ed.ac.uk/information-services/computing/comms-and-collab/office365/onedrive-for-business

The benefits of using an online cloud service is that you can upload or sync any document from your local computer to OneDrive and this will then be available for you to use in any other computer. You can download OneDrive on your own personal computer and use it to sync documents from your computer to your OneDrive space. Now, I do not recommend using OneDrive as your only storage place.

OneDrive is particularly useful for sharing documents with others safely (rather than through email attachments); and also, documents saved in OneDrive are easily recoverable or rolled back to their previous versions.

5.2. BACKING-UP YOUR WORK

Now, the benefit of using your M: drive is that this is backed up regularly and automatically by the University. If you decide not to use it, then I'd strongly encourage you find an alternative method to back up your work. Now what does this mean?

There are 'two levels' of backing up your work. The fist is versioning, which is saving your progress, and storing new versions of your files as you work on them. You should do this regardless of where you save your files. I recommend that, while you work, you should keep copies of various drafts along the way. Make sure you have a strategy so you are certain which file is the current draft. Old drafts should be kept as a backup and not worked on again

The second is about keeping a back-up copy of your data and documents. This is to ensure that, if something happens (e.g. you lose your laptop or your computer breaks – it happens!), you don't lose all your hard work and you can recover your data.

Backing up your work in a different place: If you decide not to use your M: drive, then the University recommends that you:

- Make at least 3 copies of your work.

- Save at least two of these on different types of media storage (for example, a laptop hard drive, and an external hard drive).

- Keep these in different locations, with at least one of them offsite.

- Check that they work regularly.

Design a strategy for backing up your work and stick to it. Decide which of these storages spaces you will use every day, and which will be your back up. Then decide how often you will back up your files - daily, weekly, monthly?

5.3. FILE ORGANISATION

Since you will have a lot of different types of data and files — raw data, anonymised data, notes, literature, sections of writing, and different versions of those — it is important that you plan in advance a strategy for how you will organise it all. I would suggest that you use a hierarchical structure of folders and files within them.

Think carefully and methodically about naming your files. Avoid vague names, and — probably the most common and dreadful file names: "chapter 3 final FINAL" or "chapter 3 THIS ONE". Include version control in your file naming strategy (e.g. v1.0, v1.1; or the date of the version).

A file name can be a trail of information that can enable you to find files quicker. For example, let's take file name IP02R0120190321:

I = interview (type of data)

P (n) = participant ID (participant number 2)

R (n) = researcher ID (researcher 01)

Date of interview in the form of YYYYMMDD

Your file naming strategy does not have to be exactly like this, but you should design some kind of naming strategy that enables you to establish a good directory structure, and more generally, good data practices.

5.4. FINDING SECONDARY DATA: DATA REPOSITORIES

If you are interested in conducting a dissertation project using secondary data and secondary analysis (that is, reusing data produced by another project), the place to look for data are data repositories. These are:

- Archive and preserve data sets.

- Able to offer and manage access to those data sets (most of the time this is done online).

- Provide other services, such as training and learning support; maintain data catalogues; and seek to improve data interoperability and standardisation to enable data reuse.

5.4.1. EXAMPLES OF DATA REPOSITORIES

https://www.cessda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide/7.-Discover/Data-repositories-as-data-resources

The link above directs you to the CESSDA Archives guide. CESSDA stands for Consortium of European Social Science Data Archives and it seeks to be a 'one stop shop' for European data (social science data, mostly). In this link you will be able to access the CESSDA archives, but also other European and international archives. It also contains a very comprehensive guide to finding and dealing with secondary data in data repositories.

The UK Data Service (UKDS) is the UK's largest collection of social, economic and population data. The data collection includes major UK and cross-national surveys, including many government sponsored surveys and longitudinal studies and several cohort studies following individuals born in 1958, 1970 and 2000. There is data from the UK Census from 1971 to 2011 and qualitative data collections containing in-depth interview transcripts, diaries, anthropological field notes, etc. You can access it here:

https://ukdataservice.ac.uk

In addition to those, you can also search Google Datasets:

https://toolbox.google.com/datasetsearch

If you are conducting research about the natural environment, the Natural Environment Research Council (NERC) data repository is also a good place to explore. It includes links to data repositories for British Oceanographic data, Centre for Environmental data, Environmental Information data, National Geoscience data, UK Polar data, and the Archaeology Data Service:

https://nerc.ukri.org/research/sites/data/

For those of you conducting research in the natural sciences, Nature provide a list of recommended data repositories for the fields of Biological sciences; Health sciences; Chemistry and Chemical biology; Earth, Environmental and Space sciences; Physics; and Materials Science:

https://www.nature.com/sdata/policies/repositories

5.5. ETHICAL CONSIDERATIONS

5.5.1. OBTAINING INFORMED CONSENT FROM PARTICIPANTS

Before collecting any primary data from participants, it is crucial to gain informed consent from them. This consent should be:

- freely given,

- informed,

- unambiguous,

- specific,

- and involve a clear affirmative action.

This means that, all consent forms should be accompanied with a comprehensive and clear information sheet including:

a) a summary of your research project;

b) what taking part in your research will involve;

c) how the data that you collect will be used, analysed and by whom;

d) how the data will be anonymised (if applicable);

e) how the data will be stored and disposed of/or shared (if applicable) after the project is over.

If appropriate, you should also include how the results of your research will be used — for example, if you plan on presenting at any academic conferences.

If you would like some examples, the UK Data Service has the resource below for students undertaking their dissertations, and this includes a good example of a consent form:

https://www.ukdataservice.ac.uk/media/622144/dissertations_and_their_data_promoting_research_integrity.pdf

5.5.2. WORKING WITH SENSITIVE DATA

THE UNIVERSITY
of EDINBURGH

is information SERVICES

What is 'sensitive data'?

Sensitive data is:

- Data concerning human participants. This kind of data is often referred to as 'personal data'. Personal data identifies individuals, either directly (e.g. a name) or indirectly (when combined with other information to identify specific individuals). Some categories of personal data are more sensitive than others (e.g. medical data) and they need to be dealt with even more carefully.

- Data relating to species of plants or animals. This is data that includes information on rare or endangered species, or other conservation activities.

- Commercially sensitive data. This is data where disclosure could cause economic harm or prejudice the interests of any person. This includes information such as references to ongoing negotiations, trade secrets, or data generated as part of a commercial funding agreement.

- Data that poses a threat to others. This is information which, if made available, would pose a threat to national security or would have a negative public impact.

Privacy, confidentiality and disclosure

Research participants have the right for their data to remain private, and therefore any handling, storage and sharing of their data must be managed carefully to preserve the privacy of the subject. It is your responsibility, as the researcher in your dissertation project, to ensure that any sensitive data is treated confidentially at all times. Now, to ensure this, there are two (main) steps to take. The first is storing your raw data appropriately. And the second is to anonymise your data so that it is no longer sensitive (i.e. any possible identifiers of a particular individual have been removed).

The University regulations require that you store your raw data (non anonymised data) on your University managed storage (your M: drive or OneDrive) or on an encrypted laptop or external hard drive. You should never use third party cloud-based storage providers. All non-digital pieces of sensitive data should also be stored securely (e.g. in a lockable cabinet or drawer).

Anonymisation vs Pseudonymisation

It is important to mention that there is a difference between anonymising data and pseudonymisating it:

Anonymisation is the complete removal of any identifiers (by either blanking out, or by replacing them with other words). This means irreversibly preventing the identification of the individual to whom the data relate. This

individual will not even be identifiable anymore when linked with other information which is available or likely to be available.

Pseudonymisation is, essentially, the same process as anonymisation — replacing any identifiers with pseudonyms — except that in this case, there will be a key to re-identify the individuals, which will be kept separately.

Think carefully about which approach you will use, and why.

Anonymising qualitative data

- Think about what identifiers should be removed and create a plan for anonymising your data (this way you will anonymise your data methodically and in the same way. For example, will you identify any replacements that you make? And if so, how (e.g. with footnotes, brackets?).

- Avoid blanking out. Use pseudonyms or replacements.

- Avoid over-anonymising. This means removing or aggregating information to the extent that you are distorting your data or making it misleading.

- A good practice is keeping a log for any anonymisation changes that you make to your data, and keep it separate from any anonymised data files.

Anonymising quantitative data

- Remove direct identifiers (e.g. names, address, institution and photos).

- Reduce the precision or detail of a variable through aggregation (e.g. birth year instead of date of birth; occupational categories rather than job; area rather than village).

- Generalise meaning of detailed text available (e.g. occupational expertise).

- Restrict the ranges of a variable to hide outliers (e.g. income or age).

- Combine variables (e.g. create non-disclosive rural/urban variable from a place variable).

Anonymising audio-visual data

You can manipulate your audio and image files to remove any personal identifiers (e.g. voice alteration or face blurring). You can also design a plan to avoid including personal identifiers earlier in the research process, as

part of your research design. For example, if you would like to conduct and film walking interviews, you could make sure that the research participant's face does not appear in the video.

5.6. METADATA AND DOCUMENTATION

Documentation and metadata are, essentially, what anyone else would need to a) find (metadata); and b) make sense of your data (documentation).

Documentation is contextual information about your dissertation data that you are likely to produce during the course of your research, and this information will aid anyone else to reuse your data. What documentation looks like will vary greatly depending on your data; but here are some examples:

- A "clean" copy of the confidentiality and consent agreement used.

- Data collection methods and procedures. Notes about your research design and methods, notes on instruments used to collect data and analyse data; plus information on the conditions of data collection.

- Data collection tools. A copy of the questionnaire(s), prompts, and/or interview schedule(s) used in the research.

- Database schemas and data structure. Variable labels and descriptions, an outline of relationships within the dataset.

- Coding schemas. Definition of coding conventions used – including information on missing data, categories, classifications, acronyms and annotations.

- Data modifications. Specification of any weighting used, identification of derived variables and the syntax used to create them, output files, and subsequent modifications to the original data.

- Quality control measures. Details on activities undertaken to verify and clean the data, an outline of formatting applied to the data, an explanation of file naming conventions, and if needed, a statement on known problems with the data.

Keeping good track of your data documentation will enable to make the most of your data — and your dissertation — for three reasons. First, keeping track of your documentation will make writing your methodology chapter much easier. In your methodology chapter you will need to include a discussion of your research design, your methodology, and the choices you made along the way. And, in many cases, keeping a methodology diary of all of these will mean that you can write with more certainty about those, rather than trying to rely just on your

memory. Second, this will show a strong, reflexive and transparent piece of research. And this is key to any excellent dissertation. And third, documentation allows you — or anyone else — to make sense of any data that you collected. This becomes particularly important if you decide to share your dissertation data in a data repository, which is covered in the very last section of this handbook.

Because of all of this, I would strongly encourage you to keep whatever form of documentation is relevant for your dissertation project, and particularly a research diary.

Metadata, unlike documentation, is standardised data about your data. The reason why metadata exists is to allow for data preservation (if you were to save your data in a data repository); discovery for sharing; and data citation. Think of it as the equivalent of a library catalogue information for a particular book. Thus, metadata is likely to have mandatory fields (title, Principal Investigator), others recommended (language, contact information), and some optional (ownership, retention period).

There are various metadata standards, often tailored to particular types of need (archiving, librarianship) or disciplines. Metadata schemes are designed for computers to read, but human readability is not excluded. The primary metadata standard for social, behavioural, and economic research is Data Documentation Initiative (DDI), designed to describe all stages of research in the social sciences by providing definitions (semantics) for every element of the data from conceptualization, collection, processing, dissemination, analysis, archiving, and, eventually, reuse.

6. WRITING-UP: DEALING WITH VERSIONS

Once you have gathered all the data that you need to answer your Research Question — whether this is just literature, secondary data, artefacts, or primary data, the next step will be working with that data. That process will vary incredibly from dissertation to dissertation, however, generally it will involve working with your data in some form and creating new files — whether in the form of text, as sections of your dissertation main body; or other data files which you may then analyse again. For example, from your original interview transcript, you may produce other files that contain sections of that interview organised by themes; and you will then work on those secondary files. Regardless of the particularities of your own dissertation protect, it is likely that you will produce a new set of files during both your data analysis and also your writing-up processes. And, as in the previous sections of this workshop, I strongly suggest that you think about those in advance, and plan methodically how you will handle them.

7. WHAT ABOUT AFTER YOU HAND-IN?

Finally, this last section will cover what you can do with your dissertation data after you hand your dissertation in, and after you leave University. The University of Edinburgh will only store your data for a limited amount of time after you graduate. And after this time has passed both your M: drive and your cloud storage will be emptied and your data deleted. Now, if you would like to preserve your data, safely and potentially share it for it to be re-used by other students or researchers, or even to have it archived, you may want to explore the option of data repositories.

7.1. DATA REPOSITORIES

7.1.1. EDINBURGH DATASHARE

Edinburgh DataShare is a digital repository of research data produced at the University of Edinburgh, hosted by Information Services: https://datashare.is.ed.ac.uk.

The benefits of sharing your data on a data repository are:

- A big advantage of depositing your data is that they will be preserved - even for your own future use.

- The data submission process creates a permanent record, a persistent identifier (DOI), and a suggested citation, so that your work can be formally attributed when re-analysed by others.

- Your data will be discoverable through Google and other search engines to maximise visibility and impact. The service can provide you with usage statistics so you know when your data has been downloaded.

- You do not need to maintain your own website to deliver your data; once deposited, management of your data is assured by Research Data Service staff.

7.1.2. FIGSHARE

Figshare is another digital repository. This one is an open source and free digital repository; and anyone can use it. You can access it here: https://figshare.com/

7.1.3 ZENODO

Zenodo is a data repository commissioned and developed by CERN, which is the European Organisation for Nuclear Research. This is their website: https://zenodo.org

7.2. ARCHIVING AND SHARING YOUR OWN DATA

If you are interested in submitting your dissertation data to a repository, my advice is that you speak to your supervisor so that they support you and give you advice as to how to do this. Submitting to a repository involves making sure that your data has the correct documentation and metadata, so that it is easily archivable and also discoverable.

WHAT NOW?

I hope that you have found this handbook useful and you have picked up some helpful tips ahead of (or during) your dissertation journey!

Before you finish reading this handbook, I thought I'd give you a few links to more resources:

- **MANTRA: Research Data Management Training**

  https://mantra.edina.ac.uk

  This is a course that has also been created by the Research Data Service team at The University of Edinburgh. MANTRA is a free, online non-assessed course with guidelines to help you understand and reflect on how to manage the digital data you collect throughout your research. It has been crafted for the use of post-graduate students, early career researchers, and also information professionals. It covers a lot of the topics that I have covered here, but in more depth; plus many more. I have taken this course myself, I found it very useful, and I very much recommend that you have a look yourself too.

- **Institute of Academic Development courses**

  https://www.ed.ac.uk/institute-academic-development/postgraduate/taught/courses-events

- **Research Data Service website**

  https://www.ed.ac.uk/information-services/research-support/research-data-service