THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

Research Data Toolkit

Handling research data effectively, safely and legally: A guide for researchers, departments and students

Introduction and Aims of the Toolkit

This toolkit brings together key information on research ethics, data protection, research data management as well as security and travel and risk assessment.

It can be used in departments to support students, researchers and departmental staff to promote awareness and good practice relating to research data throughout the research lifecycle.

The toolkit sets out the minimum requirements for the management of data by researchers and students in LSE departments. However, if a department has additional requirements they can add these into the toolkit. For an editable copy please contact Datalibrary@lse.ac.uk

Our thanks go to CASE who developed their own Data Toolkit on which this is based.

 If you are an undergraduate or taught postgraduate, look out for this icon in the contents table which will direct you to the advice most relevant for your projects.

Contents

The Basics: Ready, Steady, Go...and Stop

READY **When you are planning your research**	- Consider whether you need to complete a **research ethics review** (p.6) - Begin thinking about your **Data Management Plan** (DMP) (p.6) - Clarify whether your data provider, research funder or research partners have specific data security requirements and whether LSE needs to sign off any formal agreements for access to the data
STEADY **Before you start collecting or using data**	- Complete and submit your ethics review (if required) and await approval - Write a DMP or review your existing DMP - Send any special agreements from data providers to Datalibrary@lse.ac.uk so that they can be checked and signed off by the Director of the Research Division - Fill in any documentation or registers required by your Department - If the research entails travel, complete the Notification to Travel form
GO **When you start research**	- Register any datasets you create, download or receive on your department's Information Asset Register - Send queries about data management, storage or security to Datalibrary@lse.ac.uk - Update your DMP and your Department with any changes to your plans
STOP **When your research project ends**	- Archive and publish and/or delete any datasets (including copies) in accordance with your DMP, project-level data security document and funder requirements. - Once data has been anonymized, archived or published, think about the secure destruction of original data.

Research Ethics

An ethics review is required for any study involving:

- Interviews, surveys, focus groups, experiments, observations of people, etc.
- User generated data (e.g. from discussion forums, social media, vlogs, blogs, comments on posts or articles).
- The collection or use of any personal data/identifiable information (e.g. names, email addresses, IP addresses, social media profiles or meta-data, visual material, etc.).[1]
- Any other information that could identify a living individual (or potentially lead to their identification). For example:
    - where information from micro datasets, if combined, could lead to the identification of individuals;
    - or where an online search for particular wording could lead to the identification of an individual;
- If findings/conclusions/publication could have damaging repercussions for any individuals (reputation, stigma, bullying) or groups with protected characteristics.
- Any other reason why the research might raise ethical issues.

You must obtain approval of your ethics review before you commence any data collection

Before completing your ethics review form, you need to think about how you will handle **informed consent** (see p.8).

How to complete your ethics review online

- Read the instructions on the research ethics submission webpage.
- Complete the ethics review form using the link for researchers (student or staff).
- Remember to:
    - reflect on the questions and answer honestly;
    - discuss any questions or concerns with your supervisor, or email research ethics.

Approval of your ethics review

There are two review/approval routes as below. Applications are automatically categorised

[1] Research that will only use data from publicly available archival records (including newspapers) does not require ethics review (unless there are other reasons why it may give rise to ethical issues).

as requiring either Departmental or Research Ethics Committee review/approval. On submission forms are sent to the relevant reviewer/approver.

You must not begin any data collection until ethics approval has been confirmed.

1) Departmental[2] review

Applications that raise few ethical issues will be automatically routed to Departmental review/approval. For students, such applications will be reviewed by the project/dissertation/thesis supervisor (or academic mentor/advisor) as appropriate.

MSc student projects which are not for dissertations, and all undergraduate projects, are exempted from Research Ethics Committee review, unless the supervisor has concerns and opts to refer the application to the Research Ethics Committee.

2) Research Ethics Committee (REC) review

Applications for projects that include any of the following elements will be categorised as requiring review/approval by the Research Ethics Committee:

- sensitive topics (which participants may find emotional or distressing).
- vulnerable groups; research that poses a risk (whether physical or emotional/psychological) to either the participant or the researcher beyond that normally encountered in their regular activities.
- Deception or the withholding of information as to the true purpose of the research.
- Where consent will not be obtained in writing (with an exception for the Anthropology department).

Amendments

If you need to make amendments to a study that has already received ethics approval, you should complete an Amendments form and send this to the research ethics team via research.ethics@lse.ac.uk (students should copy in the their project/dissertation supervisor/academic mentor).

The research ethics team will advise whether any further review of the proposed amendment is required (either by the supervisor/Department or the Research Ethics Committee as appropriate). Once approval is confirmed, the research ethics team will upload a copy of the Amendments form to the researcher's original ethics application submission online.

[2] For reasons of simplicity the term 'Departmental' is used here to encompass research Centres and Institutes too.

Research that may undergo ethics review elsewhere

If your research will undergo ethical review elsewhere (for example, if you are collaborating with another university, or if the research requires NHS Research Ethics Committee approval), complete the online ethics review form as far as screen E. The research ethics managers will confirm whether or not LSE research ethics review/approval is also required.

Further information:

Please see the research ethics webpage for further information and guidance.

Informed Consent

Informed consent is widely accepted as the cornerstone of ethical practice in research that involves human participants or personal data (including data taken from social media platforms see p. 17 and REC guidance on using data taken from the internet and social media platforms). The informed consent process should stress that participation is voluntary and can be ended at any point during the research (although you may need to qualify that data can only be withdrawn up until analysis begins and/or publication).

Your information sheet and consent forms (whether provided in written or verbal format) should include the following information about data management, use and sharing.

- What the data collected from participants will be used for.
- Why you are collecting data from a research participant.
- How you will collect and store data securely.
- Whether you use third parties to collect or process data e.g. survey companies, hired transcribers or translators.
- How data will be anonymised.
- Who else will have access to the data.
- What will happen to the data at the end of your project.
- If your research is funded and are required to archive your data you should include this in your consent forms.

Please refer to the full LSE guidance on Informed Consent. It includes sample templates that researchers can adapt.

There may be some circumstances where gaining informed consent is not possible – for instance, in the case of some anthropological field work, or some research in behavioural science - but in these cases justification will need to be provided to the Research Ethics Committee.

If your research will involve children or vulnerable groups, please also refer to the guidance on 'Research with children and other vulnerable groups' as well as the LSE Safeguarding in Research and International Activities Policy.

Data Protection Legislation

For the main overview of how Data Protection affects research see here: Data Protection and Research.

- **'Normal' versus Special Categories personal data**
  Personal data is anything that is about or identifies a living individual (even if this is in the public domain). 'Normal' personal data is things like names, addresses, email addresses, height, school attended etc. Special categories personal data includes: Race or ethnic origin; Political opinions; Religious or similar beliefs; Trade union membership; Physical or mental health; Sexual life; Biometric and genetic information; Criminal or alleged criminal offences; and Participation or alleged participation in court proceedings.

- **Lawful basis – why are you processing the data?**
  You will be processing the data for research purposes, but you still may have to declare what this lawful basis is to research funders or data providers. For normal personal data, use Article 6(1)(e) Task in the public interest as the default lawful basis. For commercial research, use Article 6(1)(f), legitimate interests. For special categories personal data, use Article 9(2)(j), Research.

- **Privacy notice – what are you telling data subjects about your research?**
  You have a requirement to tell data subjects about your research. This can be done via the Informed Consent form (which should also include your lawful basis). You are also supposed to tell people when you have received their data from another source, though there is an exemption relating to research for this. You can also link to the relevant privacy notice for research subjects or health research subjects.

- **Letters to funders and privacy impact assessments**
  Funders may require a letter from the Data Protection Officer regarding compliance with data protection legislation. Use the contact below and provide your ethics form and research proposal. Funders and/or data providers may require a privacy or data protection impact assessment looking at the personal data risks to individuals of the research project. The form is here, but contact the Data Protection Officer for help filling it out.

- **Profiling and automated processing**
  Individuals have a right not to be subject to a decision made about them via automated processing, including profiling, that has a legal effect on them, unless this relates to entering a contract, is required by law or the individual has explicitly consented to the automated processing. This could apply to research you are conducting on behalf of an organisation. If this is the case, you need to be able to explain the automated processing. The Information Commissioner's Office and the Alan Turing Institute produced guidance on how to do this. You may also need to provide a human review of the processing.

Data Protection and Contracts

For the main overview of how Data Protection affects contracts see here: Data Protection and Contracts.

Where personal data is being shared with research partners or collaborators, the right sort of contract needs to be used. Templates are available from glpd.info.rights@lse.ac.uk or secdiv.contracts@lse.ac.uk depending on:

- Who makes the decisions on the personal data?
    - Data controller – makes the decision on what the personal data will be used for.
    - Joint controllers – both decide what the personal data will be used for.
    - Data processors – do what the data controller tells them to do with personal data.
- Where is the data going?
    - UK.
    - EEA – EU plus Norway, Iceland, Liechtenstein.
    - Adequate countries – UK has accepted the EU list.
    - Any other country.
- Amount of data being shared?
    - Access granted to LSE systems.
    - Data being transferred.

Data Management Planning

What is Data Management?

Data Management is essentially about looking after the data you use or create for your research. Being proactive about it about will help you to:

- Reduce the risk of losing material (and potentially cause a damaging data breach).
- Ensure you will handle research data in line with data protection legislation.
- Ensure you will handle research data in line with any contractual data security requirements from your funder, data provider, or research partners.
- Save time during your research as you will keep well organised and documented research data.
- Make your data last longer so you and others can re-use it in future research.
- Ensure you meet any funder requirements for data management and sharing.

For a great example of why research data management matters watch this short video

Data Management Planning

A Data Management Plan is a structured document that outlines how data will be collected, used, stored securely, described and shared during and after research projects.

A Data Management Plan should be completed if:

- Your research involves humans.
- A secondary data provider requires one to be completed before giving access to their data.
- An external funder requires it during the application process or once funding has been awarded.
- You are in receipt of external funding to do research at LSE.
- You are accessing any secure data that requires a license agreement or data sharing agreement.

- You are accessing any internal LSE data.

However, we generally encourage all researchers and students to fill in a Data Management Plan as it is a useful thing to do!

You can find more information about writing a data management plan and the LSE templates you can use here. If you are planning a funding application you can and should check what your funder requires you to do for data management planning here as they often have very specific templates.

A Research Data Librarian can review completed data management plans and provide key advice which should be followed. If you have any questions about writing a data management plan or would like a one to one consultation please just write to datalibrary@lse.ac.uk.

Know Your Data

The first step in effective data management planning is to identify all of the different types of data you will be working with on your project.

- Take a wide view of your data. The Concordat on Open Research data defines many types of research materials and results as data which is a great starting point.
- What secondary data are you using, under what conditions and how will you get access to it?
- What methods will you use to collect new data? e.g. observations, surveys, interviews?
- What type of data will these methods produce? e.g. notes, photos, spreadsheets, text files, audio or video recordings?
- What formats of data will you be working with? Using or converting your data to sustainable and open data formats will ensure you can access and use your data longer. See UK Data Archive Recommended Formats for more information.
- How much data will you generate? e.g. 50 hour long recordings, 100 spreadsheets, 50 word files of transcripts and will you have enough storage?
- Are you working with confidential, restricted or public data? Please see LSE Information Classification Standards.
- What will you need to complete your research i.e. what research tools and software will be required?

Collecting and Using Data

At the data collection stage, it will be necessary to think about terms and conditions, participant consent, software, services and equipment involved in data collection.

Using Secondary Data

- It is best to begin by finding out what is already available to you. Most of the Library's datasets are already accessible from home. For further information please see Library Data & Statistics; Library Companion for Data Users; Data Library Resources Reading List.
- Ensure that you follow the Terms and Conditions of the data you use for your research, whether that is data via LSE Library or publicly accessible data, e.g. WHO, UN data or UK Data Service data.
- In some cases you may gain access to internal data which is not otherwise made available for use, e.g. from an organisation such as an NGO. It is common and advised that you set up a data agreement which lays out considerations on access, use, storage requirements. These agreements should be signed by LSE on your behalf, please send agreements to datalibrary@lse.ac.uk, so they can be checked for you.
- Carefully check if there are any costs for accessing the data e.g., you may need to pay a subscription fee to access a website or an administration cost to a data supplier.
- Do you need access to the secondary data under special conditions or having completed any special training? Some data suppliers such as the ONS require you to become an accredited researcher before you can access datasets.
- Are there any other explicit or contractual data security requirements from your data provider, research funder, research partner, that you're obligated to meet?
- Consider if there are any restrictions that may limit your freedom to re-use the data in the way you want for your project e.g. future commercialisation or dissemination.
- Are there any copyright issues relating to secondary data e.g. archival material or photographs which you will need to consider?

Using Personal Data

- Remember that personal data is much wider than people's names. It also includes addresses, postcodes, photos, voice recordings, social media handles and emails.
- Combinations of data can also disclose someone's identity e.g. gender + ethnicity + age + company name may be enough to identify somebody.
- Don't collect personal data unless it is necessary for your research and plan out data collection so that it is easy to remove identifiable information from your dataset.
- Always store personal data e.g. voice recordings in a separate location to anonymised data e.g. transcripts.
- Store consent data completely separately from any data.

Using Social Media Data

- If you are using social media data remember this is personal data and you must consider the ethical implications and processes in the same way as you would for research with humans in person.
- You should also double check the terms and conditions of the social media platform you will be using data from. See the LSE Social Media, Personal Data and Research Guidance.
- Please refer to guidance from the Research Ethics Committee on using data from the internet and social media: ethics and consent.
- Remember that you will need to think carefully about what is considered public or private when it comes to social media and when you will need to get consent from the contributors to social media platforms.
- You will also need to be open and transparent that you are undertaking the research e.g. by posting about your research in your own social media accounts.
- For further guidance and advice see Social Media: A Guide to Ethics and Decision Flow Chart for Use of Twitter Posts.
- For more information on using specialist tools to collect, analyse and store social media data see Specialist Research Tools.

Using Secure and Special Licence Data

- If you need to access secure datasets from data providers under special conditions e.g. ONS, UK Data Service you may need to take specialised training from a data provider and sign a special license. Please see the UK Data Archive Guidance on Microdata Handling and Security.
- Any special licence must be checked by our Cyber Security Team first and signed by the appropriate contact. Some licenses may also require the input of our Legal team. Please email your agreements to Datalibrary@lse.ac.uk in the first instance, so we can perform the necessary checks.
- You will need to ensure you access, store and destroy data in line with any special agreement.
- Please bear in mind that agreements to access secure data can take a long time to go through all the necessary checks. It can also take a long time for them to be processed and approved on the data supplier side.

Secure data access at LSE

LSE Library and DTS can offer a number of options for accessing secure datasets both on and off campus. Please note that it is ultimately the decision of the data supplier and LSE to decide on an appropriate access route for the dataset, as not all options will be appropriate for all applications.

SafePod

- LSE Library is part of the ESRC funded SafePod network (SPN), an independent network of safe settings (known as SafePods) that provide a secure pre-fabricated setting for researchers to access datasets from participating data centres across the UK.
- The LSE SafePod is located on the fourth floor of LSE Library in the undergraduate silent zone. It is bookable via the SPN website. Please note that the approval process for applications is done by the data centres not LSE Library.
- The Library SafePod is available 10-4 on weekdays, excluding bank holidays and University closure days.
- Where applicable priority of access will be given where research is critical to society and the economy.
- Please contact the library for more information about booking the SafePod.

Secure suite

- The secure suite has an Assured Organisational Accreditation (AOC) in place, this means it has been accredited for use by the Office of National Statistics (ONS).
- The suite is a fully enclosed, locked, controlled access room located on the fourth floor of LSE Library.
- Use of the room is granted to approved researchers by appointment only via the Library.
- Access to the room is by swipe card entry, which is set up on the LSE central access control system at the request of the Library staff involved with the day to day operation of the room.
- The PC is a stand-alone PC, which can be used on or off network with appropriate port blocks inserted as required by data suppliers.
- The secure room is available during standard library opening hours on weekdays and weekends, excluding bank holidays and University closure days.

AWS server

- As the long-term solution for high performance computing, LSE's AWS (Amazon Web Services) environment provides the same functionality while utilising the cloud-hosted AWS infrastructure and resources and adhering to the principles of secure remote access to datasets and materials.
- The AWS environment comes under the purview of DTS. However, if you would like to use it, please apply through datalibrary@lse.ac.uk.
- All applications to use the AWS server will need to be cleared via Cyber Security, who will require a copy of your data management plan as well as any data sharing agreements/ licences you have been asked to sign.
- They will assess whether the server is appropriate for usage on a case-by-case basis, please note that this process can take a long time depending on service levels.
- The AWS server is not appropriate for all dataset applications, some data can only be accessed through the Rlab terminal server or on campus.

RLAB remote server

- The RLAB maintains a terminal server for use by researchers affiliated to the

research centres CASE, CEP, CVER and STICERD.

- Please note that you will need to be an ONS accredited researcher before you can access ONS datasets.

Security Sensitive Data

- If the research topic involves accessing, using or collecting security sensitive material e.g. related to terrorism or violent extremism of any kind then you may need to arrange for special access methods e.g. on LSE campus only or via a secure server.
- Please contact Datalibrary@lse.ac.uk who can arrange meetings with the appropriate contacts at LSE.

Doing interviews & Focus Groups Online

There are several ways that you can undertake and record online interviews and focus groups securely and in-line with General Data Protection Regulations (GDPR).

- Make sure that whatever method you use you first get explicit consent from the participant to either audio record or to video the interview/group.
- Whichever tool or method you use we advise that you disable the video call feature so as not to record people's faces (unless video is required for your research purposes).
- If you plan to use any other app, software or method for undertaking or recording interviews online or remotely please check with us first as we can advise on the security and privacy settings of the app. Please write to datalibrary@lse.ac.uk.

Using MS Teams

- MS Teams is recommended for conducting and recording interviews via your PC, laptop or phone (via the MS teams app).
- You can use the calendar within MS teams to invite participants to a meeting using their email address.
- Once in the meeting you can select the three dots option on the screen and then select 'start recording'.
- Recordings are saved to MS Stream which you can access from your LSE

Microsoft 365 account.

- If you need advice about using automatic transcription in Teams please contact the data library.

Using Skype

- When you have set up your Skype call and joined the call select the three dots on the screen for more options and then click 'start recording'.
- Once you end the call the recording will be posted to your chat.
- Download and save the recording to your LSE OneDrive or H:Space or Encrypted laptop immediately and remove the recording from the Skype chat.

Using Zoom

- Sign into Zoom using SSO with your LSE account name and password (see guidance from Eden Centre on signing into Zoom).
- Ensure you set up your Zoom meeting with a Password for additional security.
- At the start of the meeting click the 'Record' option at the bottom of the screen.
- Select the 'Record on this Computer' option.
- After the meeting is finished Zoom will process the recording and it will open.
- Recordings will be stored locally in the Zoom desktop app and you can access them in the 'Recorded' tab on the left.
- You can then move the recording to LSE storage e.g. your OneDrive.
- If you need advice about using automatic transcription in Zoom please contact the data library.

Doing Surveys

It is important to choose a survey solution which stores, uses and destroys data securely and in line with GDPR.

- LSE recommends and supports Qualtrics Survey software which is available for you to request access to (usually within a day).
- Another option is Survey Monkey which has been assessed by LSE to have adequate security and data privacy measures.
- MS forms is also available for you via MS Office 365.
- Please note common services such as Google Forms do not meet GDPR and

should be avoided.

Transcription Software and Services

- You will need to check if any software or services you use meet the requirements of the GDPR. If you are unsure, please contact datalibrary@lse.ac.uk and we can arrange for the specific tool to be checked by our cyber security team.
- Think very carefully about using cloud based services for storing or collecting personal data. You need to check the terms of services carefully to find out what happens to the data, where it is stored and for how long.
- If you will be employing external researchers to do surveys or translators or transcribers you should set up a non-disclosure agreement (NDA). LSE has produced a model NDA template.
- No matter what services or software you use, remember that the data always has to be stored in the EEA or countries with equivalent levels of protection.

Audio and Video Recording

- It is best to use audio recorders or dictation devices for doing audio recordings rather than your smart phone. If you are dealing with sensitive personal data you should choose a device that offers real time encryption such as Olympus DS-7000, Olympus DS-3500, Phillips DPM 8000/00.
- If you do need to use a smart device for audio or video recording e.g. your phone or tablet you need to ensure that the device is fully encrypted, that iOS or Android is up to date (including all apps), that you only have apps downloaded from the Apple or Google Play store and that the remotely wipe data options are enabled.
- If you are using a recording app on your smart device, or laptop, you need to make sure that the files are not automatically synced to a third-party cloud service and that the files are stored in an encrypted folder or drive.
- You should always transfer recordings from any recording device to a computer or LSE storage as soon as possible and then delete the files from the recording device. If you are storing these files on your own laptop or storage device, they must be encrypted.
- Please read the full LSE guidance on Audio and Video Recording.

Content Scraping, Text and Data Mining

- If you're using a content scraper or text and data mining, please check the terms of service of the content provider you plan to use - they might have set conditions for using a content scraper or text and data mining, or might require the explicit consent of their users.
- If it is unclear from the terms of service whether the tool is permitted, or under which conditions, then get in touch with the content provider to ask for permission to use the content. The Library can help you with permissions or to find alternative sources of data. Contact Datalibrary@lse.ac.uk.
- The Pro Quest TDM Studio is a platform for text and data mining across the Library's current ProQuest subscriptions. The visualisation dashboard has a geographical user interface with pre-built visualisations that can be applied to newspaper content. The workbench dashboard allows analysis across the majority of the Library ProQuest subscriptions using R and Python. Workbenches can only be created on request, so if you would like to access the TDM studio please contact Datalibrary@lse.ac.uk.

Data Security and Storage

You can find lots of information and guidance on data security and storage at the Cyber Security & Risk webpage.

Data security is a vital component for any research project but especially those involving personal data:

- Data breaches could lead to negative publicity for you or the School, could

damage your reputation with data providers and jeopardise your current and future research.

- Data breaches can happen in a variety of ways including: phishing or malware/ransomware attacks; losing unprotected laptops or storage devices; sharing data with people who don't have the right to see it; inadequate access controls; leaving paper data unsecured in offices.
- All staff and students who are handing personal data should take the LSE Cyber Security Awareness Training before data collection begins and must comply with the LSE Information Security Policy.
- Storing and backing up your research data securely is essential for preventing loss of data. It is absolutely essential if you are collecting, using and storing personal data in your research.

The following section outlines the main requirements and guidance you need to ensure you don't mishandle personal data.

Use LSE Storage

- Every researcher and student has 1TB of cloud storage in the secure LSE. OneDrive with servers held in the EEA. It is essential that any personal data you are using or collecting in your research is stored in the EEA.
- If you are doing a collaborative project you can set up a SharePoint or Microsoft Teams site with dedicated permissions set up for both LSE and external collaborators.
- Personal or confidential data should only be stored on LSE storage solutions (e.g. One Drive for Business, SharePoint, H: space, departmental shared folders for staff).
- If you are working with personal data, using external services such as Google Drive and Dropbox are not permitted because they will not guarantee their servers are held in the EU.
- Storing your data on your H: Space or other shared drive ensures data is backed up by LSE daily.
- When storing your data on Microsoft One Drive and SharePoint, you can rely on the 'version history' feature to restore the files if necessary.

Storing Data on Your Own Devices

If at any point in your research you need to hold any personal data (completed consent forms, interview notes, survey responses etc) on your own devices (laptop, mobile phone, external hard drive) you will need to:

- Apply full drive encryption to your devices. Personal data should never be kept in an unencrypted form if it is kept on non-LSE devices or storage.
- Use strong passwords for all your logons or when you password protect documents.
- Make sure you have anti-virus software installed, updated and making regular scans. LSE can provide Anti-virus software for staff and students.
- Enable the built-in firewall options in the operating system of your laptops or personal PC.
- Regularly update the operating system of your laptop or personal PC, as well as software and security patches.
- Use services to remotely wipe data of your device if it is stolen or lost e.g. through iCloud for Macs or software such as Prey for Windows.
- You should be aware of some of the risks of holding data, especially personal data on external storage devices, most significantly risk of loss. Please see LSE Guidance on Using External Storage Devices.

Encryption

LSE has a number of documents that will help you to ensure your devices and files are encrypted if required:

- LSE Data Encryption Requirements.
- LSE Encryption Guidelines and LSE Encryption Guidelines for Students.
- LSE Encryption Matrix.
- Using 7-Zip to Encrypt and Decrypt Files.
- HSCIC data access user awareness training (lse.ac.uk).

The UK Data Archive has also produced useful videos on 'Encrypting a Hard-Drive with Bitlocker' and Encrypting a Hard-Drive with FileVault.

Securing Physical Copies of Data

- Secure any paper-based research or signed consent forms in a locked filing cabinet with the key stored in a secure place. Don't leave the papers on your desk or unlocked drawers especially if sharing an office.
- If you scan paper data or forms so they are digital then you should securely shred paper copies.
- Ensure you have access to shredders if you need them or purchase a hand-held shredder if you are travelling.
- Secure devices and hard-drives in a safe place when leaving them e.g. on fieldwork or in a shared office space.

Backing Up Data

- The 3-2-1 system is an easy one to follow: Keep 3 copies of your data; on 2 different storage types; with 1 copy being stored remotely.
- LSE back-ups data stored on your H: Space and departmental shared folder. You can request to recover any data that you have accidentally deleted, as far back as one month prior.
- You can recover data stored on OneDrive or SharePoint through the 'version history' and 'recycle bin' feature.
- There may be times when data is only stored on your own devices and you will need to make sure that this is backed up so that you don't lose research data.
- Try to identify what data you really need to keep so you don't make unnecessary back-ups, especially of data that includes personal identifying information.
- Check whether your data provider or research funder has explicit backup requirements. Some research data providers do not allow any duplicate copies of their data to be made.
- For critical data you will want to have one or more recoverable back-ups in secure locations.
- Back up to the LSE network or OneDrive as soon as you can e.g. if you have been travelling and have not had access to the internet.
- Use test files to check that your files are being backed up as expected and check

files for any signs of corruption regularly.

- Please refer to additional guidance from NCSC on backup at Backing up your data - NCSC.GOV.UK and Offline backups in an online world - NCSC.GOV.UK.

Data security when travelling

- Researchers should be aware that in some countries checks of devices, emails and social media accounts, by authorities may be likely.
- In some countries it is not permitted to travel with encrypted devices and you may be asked to provide the encryption key.
- It is advised that researchers save all their data in LSE OneDrive cloud storage and remove any data from devices (including any syncing or shortcuts) before travelling through borders and to re-download the data once they have passed through the border.
- Researchers may also want to use a dedicated device and email accounts for when they are conducting fieldwork. If social media accounts are a concern, researchers could prepare some dummy accounts before traveling.
- If your research topic might be sensitive to the authority, be aware of any phishing attempts as the state monitoring often starts with an email tricking you to implant a spyware.
- If you have any concerns about travelling with your data you should seek advice from the Health and Safety Department Health.And.Safety@lse.ac.uk and DTS dts.cyber.security.and.risk@lse.ac.uk.

General Security Tips

- Always lock or log off from your computer when you step away from your machine especially if you are working in a public or collaborative space.
- Set up your computer to auto lock after 15 minutes of inactivity.
- Apply a privacy screen to your computer if you are working in a public area.
- Control access to any buildings, rooms, cabinets and devices which hold any personal data.

- Be aware of potential phishing emails which may compromise the security of your devices and your research data.
- Be aware that if you need a device repaired and it contains personal data that giving this to a repair shop or manufacturer may cause a data breach.
- Store the research data which contains personal information separately from anonymised data e.g. anonymized data on laptop and original data on LSE H: Space.

Sharing Datasets for examination

- Some departments require students to submit copies of datasets that underpin their assessed work for examination.
- For departments that have this requirement it's important that the datasets are shared with examiners safely and securely in line with data protection legislation and in full compliance with licence agreements between LSE and commercial data suppliers.
- Failure to comply with best practice guidelines in sharing datasets for marking heightens the possibility of a data breach, which can result in the loss of access to subscription data across the School, as well as possible financial liability and reputational damage to the School.
- The recommended method for sharing datasets for marking purposes is to submit them to Moodle.

Submitting datasets to Moodle

- Moodle is hosted on LSE's private cloud. In other words, it's infrastructure as a service, and the data would not be leaving LSE.

Privacy

- A Moodle course can be made private so that there is no self-enrolment option for other students. This can be done through the 'Enrolment methods' area of a course.
- If the course is limited to manual enrolments only then it would just be the enrolled students/markers and the Eden.Digital team that have access to the course.

Adding external markers

- For markers who are external and will not be provided with LSE credentials there is still the ability to enrol them to a Moodle course page. They will need an LSE Public Account for this.

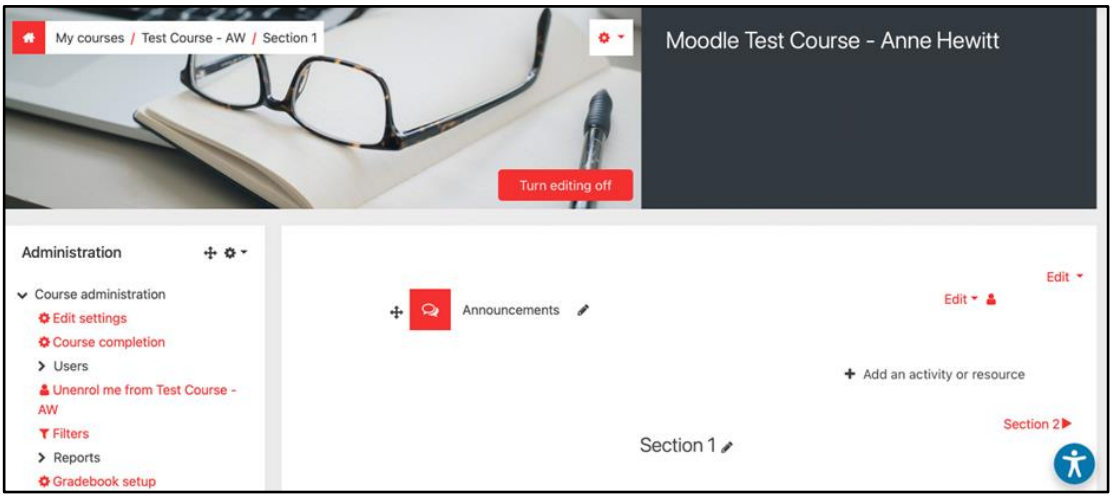
Depositing anonymously

- When using the assignment activity on Moodle it is possible to enable 'anonymous submissions' in the settings so that markers are unable to see the student identities while marking.

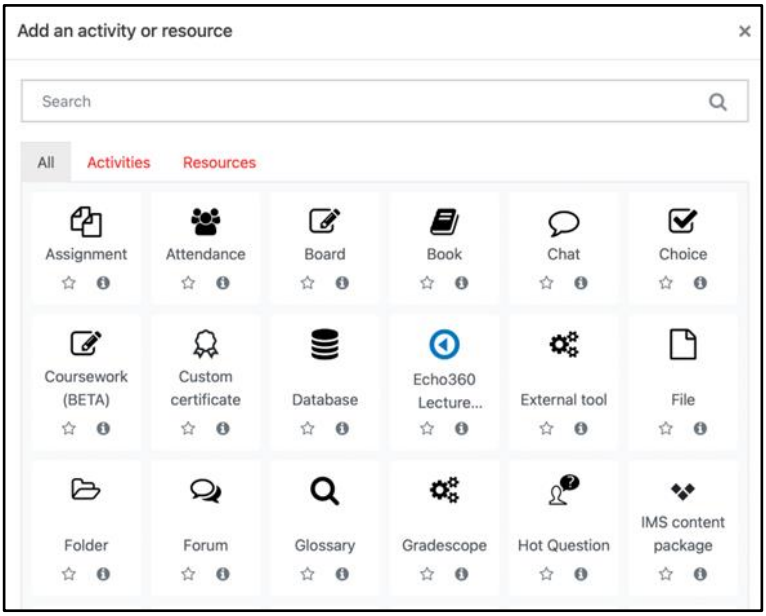
Instructions

Once you have created the private Moodle course, and added external markers if needed, you can set up a deposit for dissertation datasets.

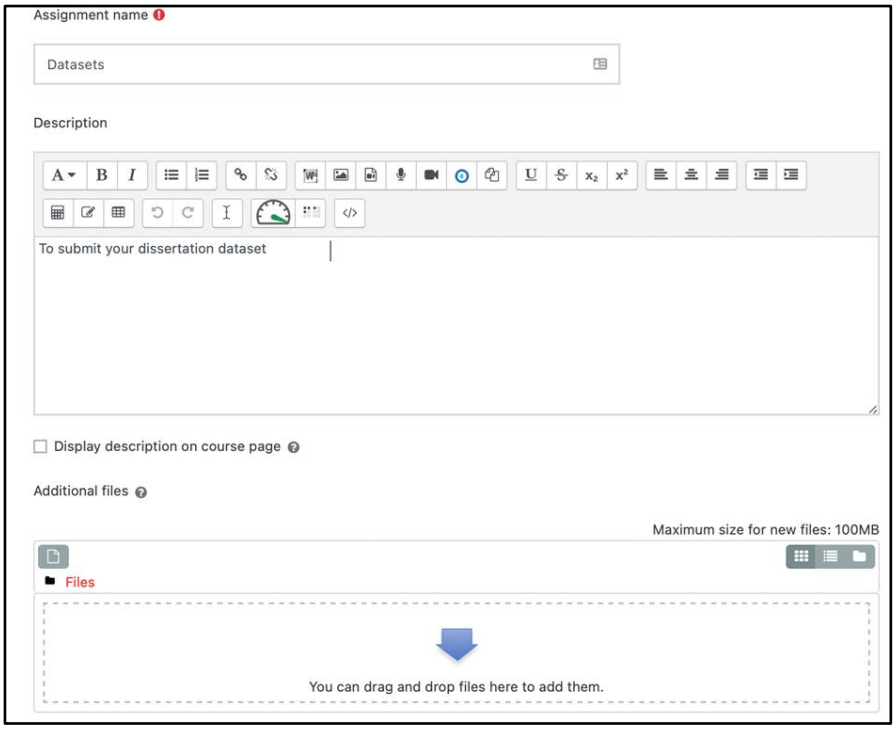
Click on **Add an activity or resource**



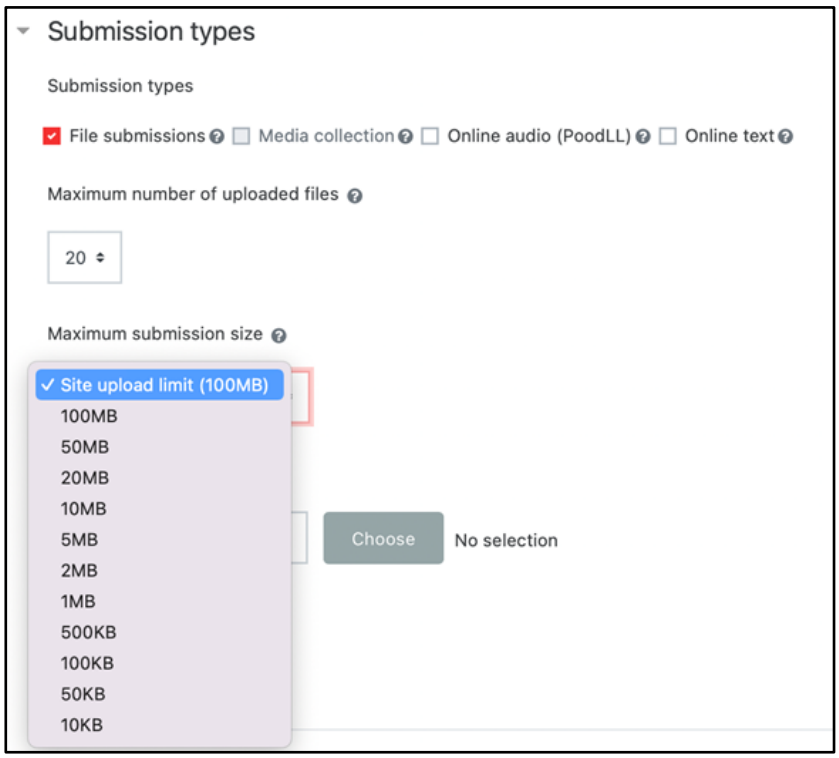
Under All, click on **Assignment**



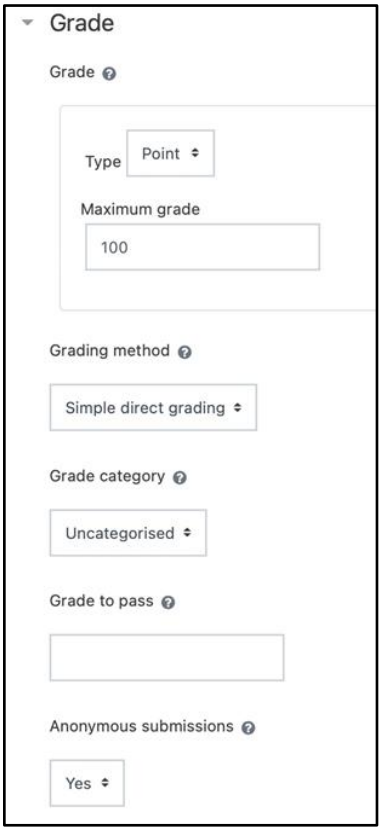
You can give the deposit a name and description



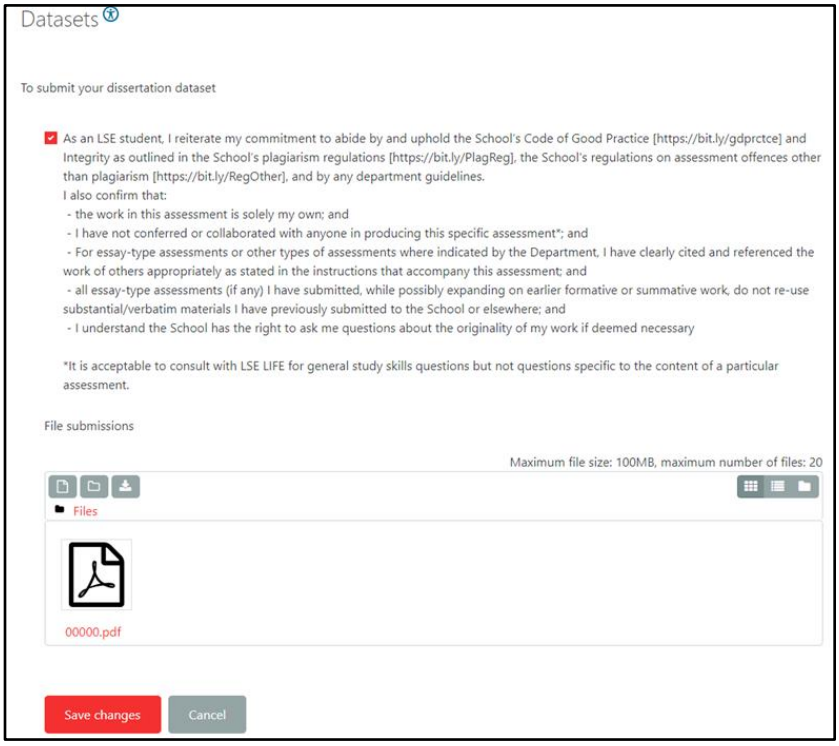
To edit submission size, scroll down to Submission types and click on the drop-down menu under **Maximum submission size**. You can also edit how many files can be uploaded, and what types of files can be uploaded to Moodle.



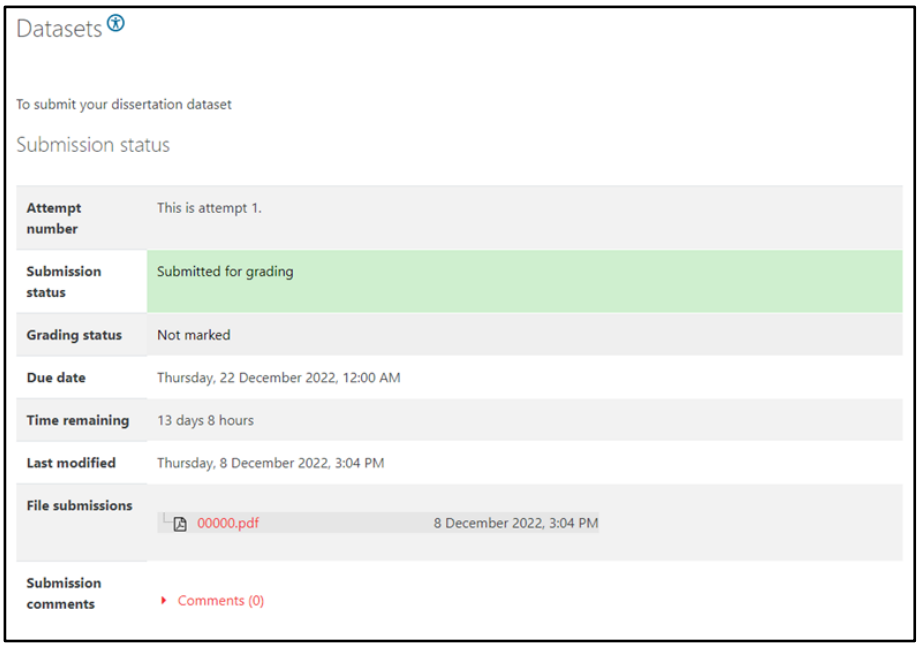
To anonymise dataset submissions, go to **Grade** and click **Yes** under **Anonymous submission**



When submitting, make sure that you are not including your name in the file name. Use your **LSE candidate number**:

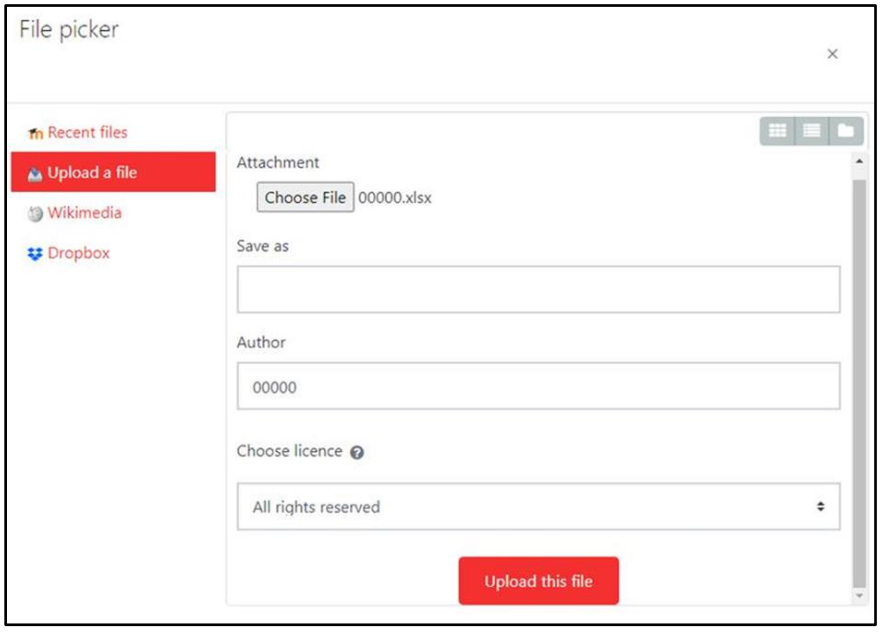


You can see the file name in the submission status:



Also, you can add additional files:

But be careful! The file picker automatically puts your name in as the author:



Working Collaboratively with Data

If you are doing collaborative research you will need to think about who needs to access your data and how you provide access securely:

- Access rights should be given following the principles of 'least privilege' (no individual should access information to which they don't have legitimate right to access) and 'need to know' (the right kind of access is given to people with legitimate access requirements). More information can be found on the LSE Access Control Policy.
- Using LSE storage facilities is the safest way to share research data. Can you use shared drives or LSE OneDrive for internal projects? For collaborations with external partners have you thought of LSE SharePoint ?
- Are any external collaborators following best practice and legislation when handling personal data?
- Are your transcribers or translators aware of data protection legislation?
- Does your research survey provider collect and process the data in a way that is compliant to the data protection legislation?
- Have you written data protection and secure use of data into your data sharing agreements, collaboration contracts, service agreements? Or have you provided third parties with a non-disclosure agreement.
- If you are sharing data via a departmental shared folder or SharePoint then you should provide access via user groups or roles rather than to individuals. Review folder permissions every 6 months.
- Access to individual files can also be controlled e.g. by password protection, read only access, read and write access.
- Email is not a very safe way to share data. It can be sent to the wrong person or can be forwarded on to unintended people.
- Never send unencrypted personal data via email.
- Consider using LSE's filedrop service if you need to send data out of the organisation.
- If you are sharing physical copies of data which include personal information you should NOT disclose the classification marking on the envelope and use a courier service. Seek advice from the LSE Records Manager if you are in any doubt.

Organising and Describing Data

Organising Data

Sensible file names and well-organised folders will make it easy to find and keep track of your data files

- Choose a consistent file naming system before you start working or collecting data.
- Don't include personal identifiers in your file names or folder structures.
- Create a folder structure that fits your research project and means you can easily locate your data. Think about which of your research resources should be grouped together.
- What are you most likely to look for first? e.g. date of interview, region, type of methodology.
- Define clear and consistent variable names carefully so they can be identified, understood and have meaning.
- Use versioning to mark-up minor and major changes to data so that you can check back on previous versions of data if necessary. When a dataset is finalised mark it as 'final'.
- Use a coversheet or template for transcriptions so that they are consistent across your project and include speaker tags to identify the question and response. Templates are available from the UK Data Archive.

For more information see the Organising Research section of the LSE Research Data Management pages.

Describing Data

It is important to think about what documentation and descriptions of your data you need to keep or create that will provide context and meaning for your data for yourself and for others.

- Study level documentation and description provides broader information on your

research project including: project aims and history; people involved; funders; data collection methods used; geographic region studied.

- Data level documentation and description provides information on a particular set of data and can include: data lists, blank copies of questionnaires, details of coding conventions, categories, acronyms and annotations, weighting specifications, identification of derived variables, modifications to original data, readme files, explanation of file naming conventions.
- It is also important to keep copies of administrative documentation e.g. blank copies of your consent forms and information sheets.

For more information see the UK Data Archive's advice on documenting data

Publishing and Archiving Data

It is becoming more and more common for researchers to publish the data that underpins their publications or research projects. Researchers may be required to do this by their publisher or their research funder. The publishing and preservation of well organised and documented research data is invaluable to advancing research.

Adding your research data to a recognized repository has many benefits

- Your data will be stored and preserved for the long-term.
- Your data will be discoverable through catalogues and other services such as Google Scholar.
- You will receive a DOI and citation for your data giving you credit if another researcher uses your data.
- You will have a choice of access conditions and licenses for your research data.
- Most data repositories are free to use.

Publishing Your Data Ethically and Safely

If your research has involved human participants you can still publish your data but you have a duty to meet ethical duties and data protection legislation.
Considerations and options will include:

- Have you anonymised your data sufficiently? Do you need to do Statistical Disclosure Control before publishing outputs or data?
- How will you license your research data for re-use by others? Does your funder require a particular license? Does your repository have a standard license agreement?
- Does your data need to be embargoed before sharing?
- Will you make your data available on request only basis so that you can assess each request for use?
- Will you restrict parts of your dataset?
- If your data cannot be shared, can you still share contextual information?
- If you didn't collect consent for publishing and preserving your datasets, can you re-contact participants if possible and appropriate? Or can you guarantee that your data can be published with full anonymization?

If you have any concerns about publishing your data, need to know what is required by your funder or want to explore options for publishing your data please contact datalibrary@lse.ac.uk.

Anonymising and Pseudonymising Data

Anonymisation is the process of turning data into a form which no longer identifies individuals and where re-identification is not likely to take place. Pseudonymisation is when you remove enough identifiers that whoever you are sharing with cannot identify the individuals.

- Personal data should be anonymised whenever possible and as early as possible in your research project.
- Can you replace identifiers or characteristics with pseudonyms?
- Can your data be aggregated from smaller to wider units e.g. for ages, income bands?
- Can you apply restrictions on upper and lower ranges of variables?
- Can you remove a variable without compromising the value of your data?
- Mark any data which has been replaced or removed.
- Remember to check any contextual information or combinations of information which may also identify participants.
- You should also keep in mind whether there is other publicly available information which when linked with your data could identify one of your research participants.
- Keep a log of anonymisation that you have undertaken.
- If you can't anonymise your research data then you need to be clear with your participants about this and tell them what you are doing to protect their identities. In this situation secure storage and sharing of data as described in this toolkit is very important.

There are a number of places you can get information about anonymization: LSE Anonymisation Web Pages; UK Data Archive Anonymisation; UK Anonymisation Network (UKAN).

Data Access Statements

Data access statements are used in published works to link to any data which underpins the publication and lay out the terms for access. According to the UKRI Open Access Policy (2021) data access statements are now mandatory for all in-scope research articles 'even where there are no data associated with the article or the data are inaccessible.'

What the statement should include:

- An explanation of how the data can be accessed, either a link to a recognised data repository (preferable) or a contact email address where the data can be requested. This email address should not be a personal contact address but rather a departmental or shared group email address.
- Any terms which restrict data access i.e. whether the data is open, safeguarded, embargoed etc. You may also wish to include a brief explanation of why data was made available on these terms.
- A copyright statement which outlines your license type and ownership of the data
- Credit for any secondary dataset you may have sourced which contributed towards the creation of your own dataset.
- In circumstances where the data has not been made available you may wish to include a statement about why it couldn't be shared i.e. 'the data underpinning this article has not been made available due to the terms agreed with our commercial collaborators.'
- If your research is funded, you should also acknowledge your funder in the statement.

Example data access statements

---

1) underlying data has been made fully available via a data repository:
Citation: Bunbury E, Mocrieff A (2022) *Experiences of data sharing amongst researchers at UK Universities*, Oxford University Press, Vol. 1, iss.1, pg.215-253, *[insert article doi]*

Received: December 1st 2021

Accepted: June 2nd 2022

Published: July 9th 2022

Data Availability Statement: The data outputs and completed questionnaires are freely available in Zenodo [inset doi link]

Funding: This work was funded by the Economic and Social Research Council (ESRC) [insert grant no.]

---

2) - underlying data has not been considered shareable:

Citation: Bunbury E, Mocrieff A (2022) *Experiences of data sharing amongst researchers at UK Universities*, Oxford University Press, Vol. 1, iss.1, pg.215-253, *[insert article doi]*

Received: December 1st 2021

Accepted: June 2nd 2022

Published: July 9th 2022

Data Availability Statement: The decision has been made that data collected are too sensitive to be freely shared, datasets are available to researchers upon request from the researcher [insert generic or admin email account]

Funding: This work was funded by the Economic and Social Research Council (ESRC) [insert grant no.]

For more information about writing data access statements please contact datalibray@lse.ac.uk

Data Retention and Destruction

Your responsibilities for your data don't end when your particular research project finishes. Whether or not you have published or archived your data at the end of your project you will need to think about data retention and destruction of any data which includes personal information. LSE provides general guidance in the Retention Schedule.

- Don't store personal or confidential data for longer than is necessary.
- As a general rule researchers are normally advised to keep research data for 10 years from the end of a research project or the point of publication of your findings for potential peer review purposes.
- If you have deposited your data in a data repository and have no further re-use for the data it can be removed from local and LSE storage.
- If you have original data which you have subsequently anonymised and you no longer need the original it is recommended that you destroy originals so as to reduce the amount of personal data you hold.
- When you do destroy data remember that normal file deletion doesn't truly erase the data and only creates another copy in the Recycle Bin. Use an 'eraser tool' to ensure data is permanently removed from your devices.
- Any paper copies that contain personal data should be shredded or LSE Porters can arrange for the secure disposal of sensitive printed material estates.porters@lse.a.c.uk.
- CDs or DVDs which contain any personal data can and should be shredded when the files are no longer needed.
- Don't forget to overwrite backup copies of data that you also have.
- If you are unsure when to destroy your data you should seek advice from any data managers in your department or the LSE Records Manager.

Travel, Fieldwork and Risk Assessment

Introduction

- The collection or use of data will often involve overseas travel. The LSE has a duty under the Health and Safety at Work Act 1974 to prevent harm to staff students travelling on School business.
- The School's travel usually falls into two categories, simple and complex travel. However, there are grey areas and Health.and.Safety@lse.ac.uk will be able to advise on these.
- Simple travel is largely brief and straightforward (e.g. attending a conference or meetings) and to low risk countries. The associated risks do not involve ethical considerations, research sensitivities or data collection methodologies.
- Complex travel may involve a prolonged period overseas, or where factors such as location, ethics, research design or the vulnerability of the traveller to unwelcome attention from the authorities or others in the host location give rise to additional risks. Crime rates, political instability, medical and other infrastructure will also be taken into consideration.

Approval Process

It is the responsibility of line managers and supervisors to approve travel in the first instance.

Simple Travel

- Simple travel should always be discussed with your staff and students. You should ensure that your staff or student has completed the Notification to Travel Form so that the Health and Safety Team are aware of the trip.
- On some occasions the traveller may be asked to complete a risk assessment for simple travel if the destination is higher risk.
- The risk assessment will need to be signed off by the line manager or supervisor.

Complex Travel

- All complex travel must be discussed in advance of travel. In all cases Risk

Assessments for complex travel will be approved by line managers and supervisors prior to being submitted to the H&S Team. Risk assessments are then sent to our external travel risk advisors for comment/advice.

- In some circumstances when the risks are extremely high the Health & Safety Team or the travel risk consultants will recommend that the proposed trip should be evaluated by the Travel Advisory Board.
- The Travel Advisory Board (TAB) will be convened where a proposed trip is to a heightened risk destination, where the nature or methodology of the research puts the traveller at greater risk or both.
- The TAB is made up of key stakeholders including academics and the relevant Head of Department or Service Leader. The TAB Reviews existing risk assessments completed by travellers, and explores further options for minimising risks.
- The Board escalates proposed travel to the School's Management Committee (SMC) when it is determined that the risks may exceed the School's appetite to accept them. In these cases SMC will make the final decision to approve travel or not.
- At any stage of the approval process the traveller and the supervisor or line manager may be asked to provide additional information regarding the trip or be asked to consider altering aspects of the travel proposal such as research methodology, location etc.
- This process can take several weeks so if a complex trip is planned please notify the H&S Team at the earliest opportunity and at least 3 months before travel.
- Please note that approval to travel can be withdrawn at any time if the safety and security situation deteriorates.

Risk Assessment

- Where the Foreign and Commonwealth Office (FCO) has a warning against non-essential travel, or the School's insurer classes the country as a high or extreme threat destination, staff and students are required to complete a Risk Identification Form and an Overseas Travel Risk Assessment. However, a risk assessment will usually be required for all research trips.
- Travel to low risk destinations for meetings or to attend conferences will not normally require a risk assessment.
- The H&S Team will inform travellers of the need to complete a risk identification form or a risk assessment once the notification to travel form

has been submitted.

- The risk identification form is used to inform the traveller and supervisor/line manager of the risk factor levels for the trip. Reasons should be given for the risk level selected.
- The risk assessment is in-depth assessment of the specific risks that a traveller may face.

Further information can be found on the Overseas Travel web pages or by contacting Health.And.Safety@lse.ac.u

Freedom of Information Requests

Under the freedom of Information Act (2005) you have the right to see recorded information held by public authorities. Freedom of Information requests (FOI) can be used as an information gathering tool for research, but it can have issues:

- For starters, it only covers public authorities.
- Try to identify who to send the request too and prepare the ground with an initial email about your research. If you are not sending to the FoI Officer, copy them in as they are more likely to track when your response is due and can chase for you.
- Think about your questions carefully. Will the question generate the responses you are asking for? Could they be misinterpreted?
- Think about the number of your questions. The more you are asking for, the more you are likely to breach the cost limit set in Section 12 of the FoI Act.
- If your questions are around the following, they are more likely to be refused as exempt. Consider if a higher or more general layer of information would work for you:
    - Personal data – the more identifiable, the more likely it will be exempt.
    - Commercially sensitive data – the test here is harm to the ability to compete. Overall figures may be fine, but the specifics about pricing will be more commercially valuable.
- Be prepared to compromise on time. Yes, you are due a response within 20 working days, but if you are happy to wait a little longer so you get your data, indicate this to the public authority.
- Be prepared that not all public authorities will reply. Yes, they are legally obliged but it may be easier to accept a lack of response rather than go through internal review, an Information Commissioner's Office complaint and on to the tribunals and courts. This will depend on how quickly you need the data.

Appendix 1: Research scenarios

This Appendix includes various research scenarios which involve the use of personal data and how this should be handled in each case.

1. **Mixed-methods primary research with ESRC Funding (John).**
2. **Secondary research with semi-confidentialised microdata (Juliet).**
3. **Secondary research with CENSUS microdata in the Safe Room (Kalinda).**
4. **PGT Research Project using Social Media (Ahmed).**
5. **PGT Research project involving original data collection from research participants (Sam).**
6. **Data Management on a large collaborative research project (Abiola).**

Scenario 1: Mixed-methods primary research with ESRC funding

John is working on an ESRC funded grant at LSE which involves mixed-methods primary research in London on the needle fixation of heroin drug users and the related opportunity of supplying injectable methadone in drug treatment centres. His design includes interviews with drug users in drug treatment centres, and a survey of the staff. John also intends to publish his data.

Ethics Review	John completes a Research Ethics Review form which raises a number of ethics issues: - Vulnerable participants. - Sensitive topics. - Physical security risk. He describes the appropriate safeguards he will put in place. He submits the review form to the REC. The application is approved.
Informed Consent	Informed consent must be obtained from both interview and survey participants Interviewees are provided with an information sheet and consent forms which describe - what the research is about. - how data will be stored, preserved and used in the long-term. - how confidentiality of participants will be maintained e.g. through anonymisation. John checks that any potential interviewees have understood the information sheet and are happy to be interviewed before he asks them to sign the consent form. Survey participants are informed about the research in the introductory page of the survey and click that they have understood and are happy to take part.
Data Storage and Security Working from home on a laptop Paper based data	-Before doing anything else, John ensures he has taken the Cyber Security Awareness course. - John keeps his qualitative sensitive data in the school's shared drive, which has appropriate folder permission. - He encrypts the data with 256-bits AES and according to LSE encryption guidelines. - He also set up the automatic screen lock of less than 5 minutes. - Anti-virus and the firewall are switched on and updated. - He connects to a private internet connection (https vs. http).

Keeping track	- Signed consent forms are locked away, with the key stored in a secure place. - John plans to destroy forms with a shredder when they no longer are in use. He keeps a master file of data to ensure its authenticity and updates it whenever major changes have been made. In addition, he also maintains old master files in case later ones contain errors.
Documenting	As John is funded he will need to deposit data in the UK Data Archive so he checks the full advice on documenting qualitative data so that he and other researchers can understand his data. - As John has mainly used NVIVO for his qualitative analysis he refers to and uses the UK Data Archive guidelines on managing and documenting data using NVIVO9. - He makes sure includes variable names, labels and descriptions and explains any codes and classification schemes used and gives reasons for missing values. He checks the full advice from the UK Data Archive on documenting quantitative data.
Organising	John already has good practice in organising data in a systematic and logical way to make it easier for himself but:
File names	John includes information about content, date, and status information in file names, ensuring that they don't contain any disclosive information about research participants.
File Structure	He organises project files in folders according to research activity (interviews/survey), data type (images, text, database), and kind of material (publication, documentation).
File Formats	John checks the UK Data Service Guidance on Preparing Data to prepare his data for deposit and then uploads his data using the self-deposit services Reshare.
Publishing	
Funder Compliance	As John is an ESRC funded researcher he has a requirement to archive, share and publish his data within 3 months of the end of his project. The ESRC recommends the UK Data Archive.
Preparing for Deposit	John checks the UK Service Data Guidance on Preparing Data to prepare his data for deposit and then uploads his data using the self-deposit services Reshare.
Citing Data	Once published by the UK Data Archive the data will get a DOI and a citation which John can include in his own publications and

	which other researchers will cite if they use John's data.
Destruction	John is aware he should only keep the identifying data from his interviewees for as long as is necessary. As part of Johns' data management plan he plans to permanently delete the identifying data once his research output is published.

Scenario 2: Research using (semi)-confidential microdata from EUROSTAT

Juliet is a researcher and Principal Investigator on a collaborative project with other research organisations. The project focuses on the implications of intra-household sharing of resources on poverty and inequality measurement across Europe using microdata from Eurostat.

Applying for data	Juliet knows that for accessing microdata there is a special application process. - She checks with LSE Library and finds she needs to complete the self-study material for using Eurostat microdata. - Juliet reads How to Apply for Microdata? the Eurostat Terms of Use and Guidelines for Publication. - To apply for access to Eurostat data LSE has to be a recognised research entity and Juliet discovers this is the case. As the project is collaborative however she also needs to check that all research partners are also recognised as a research entity and notes down the numbers. Juliet will also need written confirmation from her collaborator institutions that the researchers involved are either staff or research students. - Juliet will need to name all the people who will have access to the data in the application as well as the data manager to whom the data will be sent. - In the application she explains the purposes of the research and the dataset she is interested in. - Juliet needs to state in the application where the data will be stored securely and refers to the Library's recommended answers. However Juliet will also need to find out how the data can be securely stored at her partner institutions. - Once Juliet's application has been provisionally approved she needs to provide a printed copy to the Library so that it can be signed by the approved institutional contact as well as by her. Juliet emails datalibrary@lse.ac.uk to arrange this.
Access and Security	- The dataset Juliet needs to use is partially confidentialised and has had special statistical disclosure methods applied to it. This means that access to the encrypted files for download can

	be provided via Eurostat's server. - For some microdata sets Juliet may have been asked to access the data at a 'Safe Centre' or 'Secure Room'. - The encrypted data is downloaded from the server by the data manager and uploaded to the access location. Juliet will need to follow the special access instructions given to her by the data manager. - Only Juliet will be able to access the data at a specific location on the LSE servers for a given period. - Juliet is aware that she mustn't take a snapshot of the data or otherwise keep any duplicate copy of the data. - Juliet might be able to access some datasets remotely from her own device, in this case she makes sure her device is applied with essential security safeguards. - Juliet takes the Cyber Security Awareness course so that she is up to date with the latest cyber security threats.
Publishing	- Juliet knows that she needs to check the specific Eurostat guidelines for publication and follow them e.g. she needs to inform them of her publications so that they can be listed in their database.
Data Disposal	- Access to Eurostat microdata is only valid for the period specified in the research proposal. - At the end of the specified period the designated data manager must destroy both the original files sent by Eurostat and any derived files that use confidential data. - Juliet checks with the Library or Department Data Manager on secure methods of disposal. - Juliet, as the principal investigator, and the data manager, must provide written confirmation to Eurostat that the data has been securely destroyed.

Scenario 3: Research using CENSUS microdata from the ONS

Kalinda is a researcher working on a project focusing on multidimensional poverty and disadvantage among children in the UK. Her work includes secondary analysis of a number of available datasets and she now wants to use microdata from the 2011 Census for England and Wales. The project is already funded by the Nuffield Foundation and Research Ethics Review as well as the Data Management plan are in place.

Applying for Data	- Kalinda's first step is to look at the UK Data Service, where she learns that there are different datasets available for England and Wales with varying degree of security (public, safeguarded, and secure).
Safeguarded Data	- She first tries to see whether the safeguarded individual data would be sufficient for her research questions as per UK Data Service advice and because it will be more straightforward and quicker for her to access it. - Kalinda finds the 2011 Census Microdata Individual Safeguarded Sample (Regional)) dataset and checks conditions for accessing it, which are: being a UK research; registering with the UK Data Service; Agreeing to the End User Licence conditions. - She is satisfied she can meet them and after accepting the licence conditions online she is able to download the dataset.
Secure / Special Licence Data	- However, after some initial exploration of the data, Kalinda decides she needs to apply for the secure version of the data as the safeguarded version is not sufficient for her needs. - Access to secure microdata like this is only available for Office of National Statistics (ONS) approved researchers if they meet a selection or criteria. - So Kalinda and her collaborators on the project first fill in the online application forms and return them by email to the ONS and a request form for the microdata she needs. The form outlines the proposed research and highlights what variables that are needed and why.
Access and Security	- Kalinda takes the Cyber Security Awareness course so she is up to date with the latest cyber security threats.

	- Kalinda is expected to sign special agreements relating to access of the dataset she needs which outline the conditions of use for the data. - Kalinda and her colleagues also need to sign confidentiality agreements. - Before accessing this type of data Kalinda, like other researchers, needs to successfully complete the Safe Use of Research Data Environments(SURE) training which is run by the UK Data Service. - Once Kalinda has passed the test set at the training she is allowed to access the data. - She checks the secure microdata user guide before she starts. - Kalinda now needs to decide how she would like to access the data. The data can either be accessed via a secure room on campus, from her LSE campus PC (provided it is in a locked, secured office) or from home using an institutional laptop issued to her by LSE. - Kalinda will need to consult with the data library and the Rlab in making a decision on the best way to access the data. - Once she has made her decision, Kalinda will also need to submit a data management plan to the data library which outlines which dataset she wants to access and how she plans to access it. - Kalinda will notify ONS of her decision and they will supply her with an accredited researcher form. Kalinda will need to fill this out with details of the address she will be accessing the data from as well as the MAC and IP address of the device she will be using. - Kalinda knows that under no circumstances can raw data be removed from the ONS server. - If at any point Kalinda wishes to change her method of access she will need to notify the ONS. They will ask her to update her Accredited Researcher form, the updated form will need to be checked by the Rlab and the data library before being sent to the Research Division for signing.
Publishing	- Kalinda knows that her results should not allow for any individual

	to be identified and so follows statistical disclosure rules. - Any outputs or statistical tables Kalinda has produced can only be sent to other researchers once they have been checked by staff of the secure room who use statistical disclosure control and if required contact the data provider. - Kalinda is first given intermediate clearance so she can start sharing results with researchers and start writing up. However Kalinda can't yet release the results into the public domain e.g. in presentations or publications. - Kalinda will also need final clearance so submits her presentation and publications for checking in advance of when she needs to present them. She remembers to acknowledge the data provider in them also.
Data Disposal	- Kalinda knows that once her contract expires her access rights to all data will be removed. - However she wants to keep her access open in case you needs to revisit the data at a later date, for example when she gets comments back from peer-reviewers. So Kalinda asks for an extension which is granted.

Scenario 4: Ahmed – Using Social Media for Research

Ahmed is a Taught Postgraduate Student in the Department of Media and Communications. He is working on his dissertation which involves scraping twitter data and digital ethnography in a closed Facebook Group.

Ethics review & Informed Consent	Ahmed knows that he needs to consider ethical issues and terms and conditions of social media platforms before he starts to use any data for his research. - Ahmed checks the LSE Social Media, Personal Data and Research Guidance. For the Twitter data he can see he needs to use the Twitter API for developers to get the data he needs. - He also needs to inform users of his research so he tweets from his own account and puts a notice in his Twitter profile. As he will only be analysing and publishing numbers of tweets, hashtags, mentions etc. he will not need to collect individual consent from Twitter users. - For the digital ethnography using Facebook Ahmed can see that according to the terms and conditions of Facebook and for ethical reasons he will need to get consent from Facebook contributors to do his research. In his ethics review form he gives consideration to what impacts his joining the Facebook group might have on the other members of the group and their discussions. - Once he has received ethics approval, Ahmed contacts the group administrator to explain his research and asks to join the group. He then posts information about his research to the group and explains that if he uses any quotes from group members he will ask for individual consent and anonymise them.
Storage and Security	- Ahmed takes the Cyber Security Awareness so he is up to date on the latest cyber security threats. - Ahmed will be collecting the twitter data as .csv files for analysis. Once he has the data from the API he removes any personal identifiers e.g. twitter handle from the dataset as he doesn't need this and saves only anonymised data.

	- He will collect the Facebook data in a number of ways (screen shots, text entry into a spreadsheet etc.). Ahmed anonymises the data as it is entered into the spreadsheet by applying pseudonyms. He also paraphrases direct quotes as he enters them so they retain their meaning but cannot be identified. - Ahmed keeps an anonymization log separate from all other data so he can keep track of the original information he has changed. - Ahmed stores the anonymization log on his LSE H:Space, Facebook Screenshots on LSE OneDrive and anonymised data on his encrypted laptop. He also applies device protection measures to his laptop e.g. regular software updates and antivirus.
Publishing	- As Ahmed is a Taught Postgraduate Student his data will only be written up in his dissertation for examination. However he may want to publish this further so he included this in the consent forms for the data subjects he quoted directly. - Ahmed makes sure that he has anonymized all of the data sufficiently before submitting his dissertation.
Data Disposal	- Once Ahmed has received the mark for his dissertation he destroys any data which contains personal information which he no longer needs. - He uses recommended software for secure destruction of data. - As he is considering publishing from his dissertation he needs to keep his consent forms secure in a locked cabinet. He also moves his anonymization log from LSE H:Space to an encrypted hard-drive which he will keep secure at home.

Scenario 5: PGT research project that requires collecting data from research participants

Sam is a Taught Postgraduate Student in the department of Sociology. He is working on his dissertation which involves interviewing Housing Officers at Camden Council.

Ethics review & Informed Consent	Sam knows that because he is working with human participants, he will need to submit a research ethics review. Only once the ethics review has been approved will he be able to begin his research. - He reads the LSE guidance on informed consent and uses the template to draft his own participant information sheet and consent form. - Once he has recruited a pool of potential participants, he emails everyone a copy of the information sheet and consent form. He does this a few days before the interviews takes place so they will have time to read about the study and decide whether they wish to participate. Once the participants have signed the forms and sent them back to him, Sam stores them securely in LSE OneDrive.
Data Management	- Because Sam is working with human research participants, he will need to write a data management plan. - Sam takes the Cyber Security Awareness course so he is up to date on the latest cyber security threats. - He will visit the research data management webpages where he can find guidance for writing a plan and a link to DMPOnline, a service he can use to generate a plan template. - As part of writing the plan Sam will document his plans for: - collecting data using LSE approved research tools. - Ensuring participants are anonymised and outline how informed consent will be collected and documented. - storing and securing any data he collects. -the steps he will take to separate anonymised data from personal data. - Sam also ensures that he applies the recommended protection

	measures on any devices he uses on the project including tablets, laptops, phones etc.
Publishing	- As Sam is a Taught Postgraduate Student his data will only be written up in his dissertation for examination. However, he may want to publish this later, so he included this in the consent forms for the data subjects he quoted directly. - Sam makes sure that he has anonymized all of the data sufficiently before submitting his dissertation.
Data Disposal	- Once Sam has received the mark for his dissertation, he destroys any data which contains personal information which he no longer needs. - He uses recommended software for secure destruction of data. - Because Sam might decide to publish, he needs to hang on to some content that includes personal data, such as informed consent forms and anonymisation logs. Because he is graduating, he will no longer have access to secure LSE storage so he moves them into separate password encrypted folders on his laptop. If Sam later decides not to publish he deletes all personal data using recommended software.

Scenario 6: A large multi-national collaborative research project

Abiola is a researcher in the International Development Department. She is the PI on a research project with partners in Myanmar, India and Thailand. Some research partners come from large well-resourced universities but some partners are from small and under-resourced research centres.

Data Management Planning	- As the project has funding from the ESRC Abiola filled in a Data Management Plan at the application stage but now she wants to go back and review it as the project is starting. She sends it to Datalibrary@lse.ac.uk who can provide feedback. - Abiola had also sought advice during the application stage which means that her project costed for necessary equipment, software and team members for data management activities. - Abiola sets up a meeting with the Library, Cyber Security and Legal Team to kick off her project and ensure the necessary contracts and agreements, security infrastructure and data management processes are put in place for the project. - Abiola ensures data management issues are included in the regular meetings and workshops throughout the project and produces a set of data management protocols for all project partners and researchers to work from.
Setting Up Data Agreements	- Abiola needs to ensure that data sharing and IPR of data is worked out contractually between project partners. - She needs to check that the agreements ensure that project partners will store data in line with GDPR legislation. - As the project will employ external contractors e.g. researchers, transcribers or translators she needs to set up Non-Disclosure Agreements and she does this with the help of the LSE template.
Ethics reviews	Abiola prepares an ethics review form to cover the whole project. She knows she will also need to check whether any local ethics approvals or research permits are required in the countries where fieldwork will take place. She asks her research partners to confirm whether they need to obtain ethics approval from their own institutions.

Data Collection	- Abiola ensures that all research partners are following the same process for consent for collecting personal data. - Abiola assigns a team member from each partner institution as overall data manager who will be responsible for monitoring data collection methods and quality. - Abiola working with research partners to set up a clear structure for folders, file naming conventions and versioning. - Quality checks are put in place for equipment, data collected and analysis and transcription. - Abiola ensures that data collected by research partners is transferred to central LSE storage asap and at regular intervals.
Data Storage and Security	- Abiola worked with the Cyber Security Team to ensure that research partners met the LSE minimum requirements for data security and storage. - Abiola takes the Cyber Security Awareness Course so that she is up to date on the latest cyber security threats. She is aware that any other LSE researchers working on the project will also need to take the course. - She set up an LSE SharePoint site which is used to store data for the whole of the project. Folders can be set up with different permissions so that only those project members who need to see particular datasets will be given access. - Abiola ensured that all devices and equipment used for the project were set up with the necessary security measures and encryption.
Publishing and Archiving	- As Abiola's project is funded by the ESRC, anonymised data needs to be archived with the UK Data Archive. - Abiola is prepared for this because she included the archiving of anonymised data in her consent forms and information and she has organised data and well documented data. - Abiola and the project team assess the anonymised data and decide under which licence and access levels the data will be made available. - The team decides that anonymised survey data can be made available under the standard UK Data Service Licence. However anonymised interview transcripts will only be made available on controlled access, meaning Abiola will be notified if a researcher wants to use the data and she and the team can decide whether

	to grant access. - The data is deposited through the UK Data Service ReShare system.

Appendix 2: Revision Quiz

Quiz 1	What is the first step to maintain the integrity, confidentiality, and availability of your data? *(choose all possible answers and then continue)* a. Understand what the data is, where the data is located. b. Understand who should have access to the data. c. Understand the legal and contractual restrictions around the data. d. All of the other choices.
Quiz 2	The best place to store your current paper-based research data when leaving the office or overnight is *(choose one and continue to view answers):* a. On your desk. b. In an unlocked drawer. c. In a locked filing cabinet. d. In a high security bank vault.
Quiz 3	How can you protect the security of research participants' personal information in a dataset? *(choose all possible answers and then continue)* a. By anonymising/reducing the precision of data. b. By storing personal data (names, addresses, etc) separately from other data. c. By encrypting data files containing confidential information. d. By immediately destroying the dataset upon receipt.
Quiz 4	What would be a suitable way of transferring data to a colleague? *(choose all possible answers and then continue)*

	a. Via an unencrypted email attachment. b. Using dropbox without encryption. c. Using sharepoint. d. By handing it in person on a USB stick.
Quiz 5	For each of the scenarios below, think of preventative measures that could have been adopted *(choose all possible answers and then continue*): a. A researcher has her laptop stolen while on a conference trip. Vital research was on it and it was not backed up. b. A transcriber tells friends and family about the interesting interview he's been working on, giving details about the participants. c. A researcher encrypts her hard drive and forgets the passcode. d. A researcher shares his paper with a colleague for comments, but the latter contains confidential data and had not been cleared yet.
Quiz 6	How can you securely destroy data on a PC hard drive? *(choose all possible answers and then continue*) a. Delete files and reformat the hard drive. b. Remove and destroy the hard drive. c. Delete files and take the computer to the dump. d. Overwrite files using secure deletion software.
Quiz 7	How can you securely destroy data on a CD or DVD? *(choose all possible answers and then continue)* a. Put it through a CD shredder and throw the shreds away. b. Break it in half and throw it away. c. Put it in a plastic recycling bin. d. Cut it up in little pieces with nail scissors.

Quiz 8	How can you securely destroy data on paper? (*choose all possible correct answers and then continue)* a. Put it in a paper recycling bin. b. Put it through a paper shredder and throw the shreds away. c. Put it in well-soiled kitty litter and then throw it away. d. Rip it with your own hands.
Quiz 9	Data documentation is essential for (*choose all possible correct answers and then continue):* a. Me to understand my research data in 5 years time. b. A reviewer to understand the findings in my article under review. c. A researcher who uses my archived data.
Quiz 10	Organising my file names and folder structure consistently is *(choose all possible answers and then continue):* a. A waste of precious time. b. Not necessary until I finish my project. c. Important for sharing and future use of data. d. Good practice for my research project.

Appendix 3: Revision Quiz Answers

Quiz 1	Correct Answer = D - All of the above choices. You should understand what the data is since you cannot protect what you don't know; you need to understand where the data is located because it can have legal implications. You need to understand who needs access to the data because only then can you technically enforce the right access restrictions. You also have to understand the legal and contractual restrictions because these must be implemented in order to achieve compliance
Quiz 2	Correct Answer = C In a locked filing cabinet. This is the ideal solution to make sure your data is kept safe, especially if it contains confidential information. Incorrect Answer = A On your desk. This is not ideal if your data is important and or confidential. Particularly so if you are sharing the office. It could be less problematic if you are in a single lockable office and the only person with the key. Incorrect Answer = B In your drawer. This is not ideal if you are in a shared office. It may be less of a problem if you are in a single, lockable office and the only person with the key. Incorrect Answer = D In a high security bank vault. Security needs to be at the optimal level, not the maximum level. This is overly secure and inefficient.
Quiz 3	Correct Answer = A anonymising/reducing precision of data. This prevents identification of participants by masking disclosive information. Correct Answer = B by storing personal data separately.

	Connections should be minimised as much as possible to avoid identification of research participants. Correct Answer = C by encrypting data containing personal information. This is ideal because only people who know the key can access files. It takes over a quadrillion years to crack a strong encryption key. Incorrect Answer = D by destroying the dataset upon receipt The goals of advancing research and safeguarding participants should be balanced.
Quiz 4	Correct Answer = C Sharepoint. This is ideal. Make sure you receive training for it. Incorrect Answer = A Unencrypted email attachment. This is not secure, especially if the data is confidential. Emails remain on public access servers and are not safe to store or transfer data. Incorrect Answer = B Dropbox without encryption. This is not secure either. With cloud storage, files are held on servers in unspecified locations. Encrypt the data to prevent unauthorised access. Depends - D In person on a USB. This is fine if it is not confidential data or it is anonymised. Make sure you give the right data and that your colleague looks after it. If the data is confidential, encryption is required as USBs are very easily lost
Quiz 5	A) A researcher has her laptop stolen while on a conference trip. Vital research was on it and it was not backed up = Regularly backup your data, as this protects against data loss and corruption. Encrypt your devices to store data securely, especially when travelling. B) A transcriber tells friends and family about the interesting interview he's been working on, giving details about the participants = Choose your collaborators wisely. Ask them to sign non-disclosure agreements, together with an agreement to destroy all data once they have been returned. C) A researcher encrypts her hard drive and forgets the passcode = Keep a copy of the passcode in a separate location.

	D) A researcher shares his paper with a colleague for comments, contravening contractual obligations as the latter contains confidential data, and had not been cleared yet = Consider having all initial and intermediate outputs labelled as "not-cleared" for being shared with researcher not on the contract.
Quiz 6	Correct Answer = D Overwrite files using secure deletion software. Special deletion tools, such as the 'eraser' tool will purge the data. Correct Answer = B Remove and destroy the hard drive. A hammer will do. Incorrect Answer = A Delete files and reformat the hard drive. Deletion and reformatting only remove the links to where the files are kept, thus they may still be retrieved. Incorrect Answer = C Delete files and take the computer to the dump. Data could still be retrieved.
Quiz 7	Correct Answer = A Use a CD shredder and throw the shreds. Once shredded there is no way it could be pieced back together. Correct Answer = D Cut it up in little pieces with nail scissors. Failing access to a shredder. Incorrect Answer = B Break it in half and throw it away. Putting two pieces back together would hardly deter an intruder. Incorrect Answer = C Put it in a plastic recycling bin. And CDs should be recycled with metal, not plastic.
Quiz 8	Correct Answer = B Put it through a paper shredder and throw the shreds away. This is the ideal solution. Make sure the paper is satisfactorily shredded – traditional shredders that cut into strips are not safe. Incorrect Answer = A Put it in a paper recycling bin. This would not deter an intruder, and could attract curiosity. Incorrect Answer = C Put it in well-soiled kitty litter and then throw it away. As gross as this is, it would not deter a

	determined intruder. Incorrect Answer = D It takes 2 centimetres to find out a person's Social Security Number
Quiz 9	All Answers are correct.
Quiz 10	Correct Answer = C Important for sharing and future use of data. It is essential for sharing your data and make it comprehensible to both yourself and other researchers. Correct Answer = D Good practice for my research project. Indeed, make it your everyday research practice. Incorrect Answer = A A waste of precious time. It actually makes it easier and quicker to keep track of your data. Incorrect Answer = B Not necessary until I finish my project. You should organise your folders along the way; then there will be little reorganisation at the end of the project.

Get in touch with your key contacts

Name	Title	Expertise	Contact
Hannah Boroudjou and Julie Baldwin	Research data Librarians	Data management plans, secure data access, data access statements, data archiving	Datalibrary@lse.ac.uk
Lyn Grove	Senior Research Ethics Manager	Research ethics, research conduct	Research.ethics@lse.ac.uk
Rachel Maguire	Records Manager	Data protection, data governance and legislation	glpd.info.rights@lse.ac.uk
	Information Security Manager	Data storage and security, cyber security, secure data	dts.cyber.security.and.risk@lse.ac.uk

Come and meet us at our weekly drop ins!

If you have any questions at all about data management that have not been answered by the toolkit then feel free to drop by our weekly data drop-in sessions. These are online sessions which run throughout the term and give you a chance for a 1-2-1 with the Research Ethics managers, the Data Protection Officer, the Law & Copyright Librarian and a Research Data Librarian. Sessions are bookable through the LSE Life booking system please contact datalibrary@lse.ac.uk for more information.

Your Key Information Links

Policies and Procedures

Research Data Management Policy

Research Ethics Policies and Procedures

Overseas Travel Policy

Health and Safety Procedure: Risk Assessment

Data Protection Policy

Information Security Policy

Notification of Travel Form

General Guidance

Research Data Management webpages

Moodle Library Companion for Research Data Management

Information Security webpages

Data Protection and Research webpages

Research Ethics webpages

Overseas Travel webpages

Specific Guidance and Documents

Ethics

Informed Consent

Research Ethics Review Process

Research Ethics Review Form

Legal

Copyright Guidance

Copyright Advice Page

Data Controller versus Data Processor

Explaining Automated Processing

Data Protection and Contracts

Cyber Security

Information Classification Standard

Data Encryption Guidance

Device Level Security Minimum Standards

Data Processing outside EEA by Collaborators – Minimum Standards

Data Processing outside EEA by Individual Contractors – Minimum Standards

SharePoint Team Sites Provision

Data Management

Writing Your Data Management Plan

Open Data

Data Access Statements

Archiving Data with UK Data Archive

Review Schedule

Review Internal	Next review due by	Next Review Start
Every 3 years	September 2025	July 2025

Version history

Version	Date	Approved by	Notes
1	04/02/2022	HB	
2	04/02/2022	HB	
3	01/02/2023	HB	
4	02/02/2024	HB	
5	07/01/25	JB	Minor revisions

Contacts

Position	Name	Email	Notes
Research Data Librarian	Hannah Boroudjou	datalibrary@lse.ac.uk	
Research Data Librarian	Julie Baldwin	datalibrary@lse.ac.uk	

Communications and training

Will this document be publicised through Internal Communications?	Yes
Will training needs arise from this policy?	No