Research Data Management and GDPR: Do's and Don'ts

20 December 2018

The University of Edinburgh Research Data Service seeks to facilitate good practice in research data management (RDM) by the research community, in accordance with the University's RDM Policy and general best practice. You can find us at https://www.ed.ac.uk/is/research-data-service.

This is a quick guide to assist University users of our tools and support services to comply with the UK Data Protection Act, 2018 and the General Data Protection Regulation (GDPR) on which it is based.

Remember - the University Data Protection Officer (DPO) is the authority on interpreting and following the GDPR within the University. The DPO's guidance for researchers is at https://www.ed.ac.uk/records-management/guidance/research/data-protection.

RDM tools & support	DO's	DON'Ts
Before you begin		
DMPonline – a tool to help you create your own data management plan	Do write a data management plan to help you make appropriate decisions about your research project to comply with GDPR, ethics and security. Do include costs of data management services in your grant proposal so you use the most appropriate options. Do plan to share – through appropriate ethics approval, consent forms and anonymisation techniques.	Don't skip writing a plan just because your funder doesn't ask for one. Don't just copy from other research projects; use funder and university templates instead, and think things through for your own work. Don't wait till the last minute! The Research Data Support team can help customise your plan if you allow enough time.
Data Protection Impact Assessment (DPIA) - the DPO offers a university template for researchers	Do know who the data controller of your study is. Do know your legal basis for processing personal data. Research as a public task may mean some GDPR rules do not apply. If you are collecting data from human subjects anywhere, do carry out a Data Protection Impact Assessment. This helps you to assess risks in order to protect your data and observe rights of study participants.	If you are a student, don't skip the DPIA just because you are not a University staff member. Creating one will help you understand the risks of working with your data. Don't plan to collect sensitive kinds of information about people (special categories data) if you do not absolutely need it for your research purpose. This satisfies the purpose limitation principle of the GDPR, and prevents the need to apply additional safeguards.

While your work is in progress	DO's	DON'Ts
Discover and re-use data: Data Library portal	Do consider existing sources of data for your research. By analysing data that is already published, you will likely be exempt from GDPR regulation. If using personally identifying data that someone else has collected, do understand your obligations as a data processor. You may be asked to sign a legal (data use) agreement.	Don't transfer data by unsafe means (e.g. email, unencrypted portable drives) if it contains personally identifying information. Don't use data available on the internet without understanding the terms of use as well as the ethical implications of processing data about people (e.g. social media data).
Storing live research data – DataStore, OneDrive	Do make use of University facilities for researchers such as DataStore, which is backed up and managed in University data centres. Do manage permissions of shared folders so that only approved researchers can access personal data. Do encrypt folders which store personally identifying data at rest. If you are aware of the risks, do use other University provided storage solutions for research data, such as OneDrive or Office365. The University has a GDPR-compliant contract with Microsoft to provide these cloud-based services to staff and students.	Don't process research data on cloud services such as Dropbox without seeking advice. Cloud services not under contract to the University may not be stored on a UK or EU server, and may not be GDPR-compliant. Don't keep copies of data with personally identifying information on portable media unless safeguards (such as encryption) are used.
Sensitive data – Data Safe Haven or other safeguards	Do apply additional organisational and technical safeguards to personal data that are in special categories*. Processing such data risks putting individuals in danger of harm, including unlawful discrimination. Do control access permissions and use encryption, anonymisation, or pseudonymisation techniques to protect personal and sensitive data. The Data Safe Haven provides a managed storage and computing	Don't carry out research on sensitive data without taking training to ensure you know how to keep it safe. Don't let copies of sensitive data proliferate. Don't access sensitive data over an insecure network (such as public Wi-Fi).

* The legal special categories in the GDPR are: race and ethnic origin; religious or philosophical beliefs; political opinions; trade union memberships; biometric data used to identify an individual; genetic data; health data; data related to sexual preferences, sex life, and/or sexual orientation.

	platform for data requiring additional safeguards. Do apply to use the Data Safe Haven for research projects using sensitive data, especially if your data provider requires restricted access or proof of security.	
Collaborating – DataSync, SharePoint	Do ensure that any collaborators understand their obligations for controlling or processing personal data. Consider using a data use agreement, or other legal document, with collaborators. Do make use of University collaboration tools such as DataSync, for making personally identifying data available to collaborators safely. Office365 tools such as SharePoint may be another option, if managed carefully. Do anonymise or pseudonymise data before sharing with collaborators, whenever possible.	Don't make personally identifying data available to collaborators unless necessary. Don't use email attachments to transfer personally identifying data without safeguards, such as encryption. Don't retain personally identifying data on collaboration platforms longer than necessary.
Version Control Systems – Subversion, Gitlab	Do manage your documents and shared codebase appropriately. Both university systems, Subversion and Gitlab, allow you to manage permissions and versions when working with a team. If you cannot share research data because you are not the controller, do consider sharing the analysis code by itself, allowing others to re-run your code after successfully applying for access to the data.	Don't use a cloud-based open access code sharing system (like GitHub) if you cannot protect personally identifying information from public disclosure.
Towards the end of your research	DO's	DON'Ts
Recording datasets - Pure	Do create a discoverable dataset record in Pure if you cannot share data openly for reasons of data protection or disclosure control. Do record where the data reside and how potential users may apply for access.	

Sharing open data - DataShare	Do deposit your dataset in DataShare once you have anonymised it. You may need to apply additional disclosure control if there is information available that, if matched with the dataset, could identify participants. Do include a blank consent form with your deposit if participants have consented to have their personally identifying data shared.	Don't deposit unconsented personal or sensitive data in DataShare or any open access repository or website.
Storing data securely for long-term retention - DataVault	Do use the DataVault as a closed archiving option if the data cannot be anonymised and shared openly. Data will be encrypted and stored offline for the selected period of retention. Do pseudonymise your data by separating the identifiers from the substantial body of data and deposit separately. This will help to ensure that data will not be disclosed in error after you leave the University. (Your data will become a University data asset.) Do clean up your data before archiving it, and dispose of unnecessary duplicate files and early versions. Do be prepared for other researchers to request access to your data after finding the record in Edinburgh Research Explorer (Pure). Have a strategy ready for vetting requests and providing access.	Don't retain personally identifying data longer than needed for research or other purposes.
Training – online courses, scheduled and bespoke workshops	Do familiarise yourself with the basics of research data management, data protection, and information security. The Research Data Support team can advise on options for your or your team's requirements. Do take additional training if you are working with personally identifying data that falls into the legal special categories.	Don't assume that your research methods training has covered everything you need to know. Don't assume that training you took before the GDPR / UK Data Protection Act 2018 came into effect is sufficient. Don't hesitate to seek help! To contact the Research Data Support team email research-data@ed.ac.uk.