SHARING RESEARCH DATA

University of York
Information Services

Research data forms an important part of the scholarly record. The University, alongside many funders and publishers, requires that research data[1] is shared wherever possible.

However, it is widely accepted that the requirement to share data does not equate to "open" access to data in every case. In practice a researcher's ability to make data openly available for reuse by others may be limited by legal, ethical or commercial considerations. In some cases, it may not be possible to share the data at all. The principle '**as open as possible, as closed as necessary**' should be applied.

Key message: Researchers should plan for how and when their research data can be shared (if it can be shared), what might limit or prohibit data sharing, and what can be done to enable others to make correct use of it.

ANSWERS REASONS NOT TO SHARE DATA

	REASONS NOT TO SHARE DATA	REPLIES OR ARGUMENTS IN FAVOUR OF SHARING
1	My data is not of interest or use to anyone else.	It is! Researchers want to access data from all kinds of studies, methodologies and disciplines. It is very difficult to predict which data may be important for future research. Who would have thought that amateur gardener's diaries would one day provide essential data for climate change research? Your data may also be essential for teaching purposes. Sharing is not just about archiving your data but about sharing them amongst colleagues.
2	I want to publish my work before anyone else sees my data.	Data sharing will not stand in the way of you first using your data for your publications. Most research funders allow you some period of sole use, but also want timely sharing. Also remember that you have already been working with your data for some time so you undoubtedly know the data better than anyone coming to use them afresh. If you are still concerned you may embargo your data for a specific period of time (if your research is funded agree the embargo period with your funder).
3	Other researchers would not understand my data at all – or may use them for the wrong purpose.	Producing good documentation and providing contextual information for your research project should enable other researchers to correctly use and understand your data.
4	I have not got the time or money to prepare data for sharing.	It is important to plan data management early in the research data lifecycle. Data management ideally becomes an integral part of your research practice, reduces time and financial costs and greatly enhances the quality of the data for your use too.

[1] research data that substantiate research findings and/or are of acknowledged long-term value

5	If I ask my respondents for consent to share their data then they will not agree to participate in the study.	Don't assume that participants will not participate because data sharing is discussed. Talk to them – they may be less reluctant than you might think, or less concerned over data sharing! Make it clear that it is entirely their decision, whereby they can decide whether their data can be shared, independent of them participating in the research. Explain clearly what data sharing means, and why it may be important. But they are still free to consent or not. You can always explain what data archiving means in practice for their data.
6	I am doing quantitative research and the combination of my variables discloses my participant's identity.	Quantitative data can be anonymised through processes of aggregation, top coding, removing of variables, or controlled access to certain variables (i.e. postcodes)
7	The data I am using is commercially sensitive.	Your ability to share data that is commercially sensitive (e.g. data provided by a commercial partner or data collected from industry-based research participant) is likely to be defined by the terms of contract or by collaboration agreements with your commercial sponsor or partner. The need to comply with confidentiality clauses and contractual obligations would be valid justifications for withholding research data. However, you may be able to share this type of data with other researchers, subject to non-disclosure agreements. That's why it is important to negotiate with commercial partners about the sharing of research data before your research begins, particularly for publicly funded research.
8	My data have been gathered under complete assurances of confidentiality.	Why was such an assurance made? It is best to avoid unnecessary promises. Anonymisation procedures can be implemented to protect identities, but confidentially can never be completely guaranteed. You can also consider controlling access to the data.
9	I have collected audiovisual data and I cannot anonymise them, therefore I cannot share these data.	Visual data can be anonymised through blurring faces or distorting voices, but this can be time consuming and costly to carry out. It can mean losing much of the value of the data. It is better to ask for consent and share data from participants in an unanonymised form, and/or control access to the data.
10	I have used existing data obtained from a third party in my research. As I have incorporated some of this data within a database containing my own research data, it cannot be made public.	It is important to know who holds the copyright to the data you are using and to obtain the relevant permissions. You need to be aware of the licence conditions of the data you are using and what you can and cannot do with the data. Before this database can be shared, permission must be sought from the data owner. If permission is not gained, your data - with the third party data removed - may be shared. Secondary users could then recombine your data with the third party data if they were to obtain it. Cite the existing data you have used in your published research outputs.

Acknowledgement:
Louise Corti, Veerle Van den Eynden, Libby Bishop and Bethany Morgan Brett. (2011) Managing and sharing research data: training resources. Colchester: UK Data Archive. Licensed under CC BY-NC-SA / Adaptation from original.