

Responses to common objections towards sharing research data

<i>Modesty</i>	<p><i>Our data are embarrassingly bad.</i></p> <p><i>Why would anyone be interested in my data?</i></p> <p><i>It's not very interesting.</i></p> <p><i>My data are not of interest or use to anyone else.</i></p>
<p>You cannot predict how people will use your data, nor should you. Just make sure it is described well and allow users to make a judgement on use (within parameters of licences and laws).</p> <p>There are many roads for reuse:</p> <ul style="list-style-type: none"> • replicating or refining research questions, testing or other hypotheses and methods of analysis • use as teaching aids for analysis or research design • creating new datasets by combining data from multiple sources <p>A professor once said “[It’s] not a risk we make mistakes, but a certainty”. The scientific process is about refinements and corrections, not hiding. If your data is good enough to publish from, it is good enough to share.</p>	
<i>Resistance</i>	<p><i>It's my data. I don't want to share it, and that's all there is to it.</i></p> <p><i>It's not a priority and we're busy.</i></p> <p><i>I have not got time or money to prepare data for sharing.</i></p> <p><i>We can't see the benefit.</i></p>
<p>Funders, journals, academic organisations, and research institutions have formal positions and informal expectations that research data be shared to the fullest extent possible, respecting legal, ethical and commercial concerns.</p> <p>Formal policy positions often include sanctions for non-compliance that could harm your ability to publish and get future research funding.</p> <p>Much of what we recognise as Research Data Management activities are what you would do anyway. Activities like cleaning data, constructing coding schemes, undertaking anonymization for publications are what you would do regardless of data sharing requirements. Most funders recognise extra resources are needed (within reason) to create reusable data, and include statements in their policies to that effect to help support data sharing.</p> <p>As for benefits, as well as scientific benefits achievable through data sharing, it increases the visibility of your research, your research profile, and citations.</p>	
<i>Sabotage</i>	<p><i>Data sharing makes it harder to recruit participants.</i></p> <p><i>If I ask my respondents for consent to share their data then they will not agree to participate in the study.</i></p> <p><i>I don't mind making it open, but I worry someone else might object.</i></p> <p><i>If we publish this data, people might sue us.</i></p>

	<i>We don't have that data.</i>
<p>Let participants decide; don't make the decision for them.</p> <p>Make sure your information and consent process explains what data sharing is and the protections involved preventing identification and harm, like restricted access and anonymisation. If, after that, participants still do not want their data to be shared then that has to be respected.</p> <p>If you don't have the data, your funder and/or institution will want to know why not. You will also find it increasingly difficult to publish without providing information on where data can be found and how it can be accessed.</p>	
<i>Fear</i>	<p><i>People might misuse my data.</i></p> <p><i>People will misinterpret the data.</i></p> <p><i>Terrorists might use the data.</i></p> <p><i>Other researchers would not understand my data at all – or may use them for the wrong purpose.</i></p> <p><i>People will contact us to ask about stuff.</i></p> <p><i>We want people to come direct to us so we know why they want the data.</i></p> <p><i>We'll get spam.</i></p> <p><i>It's too big.</i></p> <p><i>It's too complicated.</i></p>
<p>You are not responsible for how your data is later used. Users are responsible for their own interpretations, analysis, and conclusions.</p> <p>If there are legitimate concerns about commercial usage of your research data, non-commercial usage clauses can be applied. If there are national security concerns these should be raised with your institution and funder, and suitable limits can be applied to the data.</p> <p>Size and complexity need not be a problem. Instead, think about how you are documenting and describing your research so it has contextual meaning. Ask yourself, what would I need to know to understand these data? If you don't believe other researchers would understand it, what is to say you will be able to understand it if you were to return to it in five years' time? Ideally, if you have done a good job with documentation, people reusing the data will not need to contact you to ask you questions.</p> <p>Placing your data in a reputable archive or repository can solve the question of managing requests for re-use, legitimate, illegitimate, and spam. On one side, they take away the burden of having to deal with requests. On the other, archives and repositories often have a requirement that involves potential users saying why they want to use the data and in some cases offer depositor approval, so the owner must give permission for reuse on a case by case basis.</p>	
<i>Recognition</i>	<p><i>I don't want other people taking credit for my work.</i></p> <p><i>We might want to use it in a research paper.</i></p> <p><i>I will if I can have an embargo...is 30 years OK?</i></p> <p><i>I want to publish my work before anyone else sees it.</i></p>

	<p><i>What if we want to sell access to this data? Our lawyers want to make a custom license.</i></p>
<p>People copying or taking your work without attribution is plagiarism. Data sharing involves legal and moral obligations to cite original data sources and does not form scientific misconduct.</p> <p>Data can be cited in the way publications are cited, with associated professional recognition and credit. Placing your data in a trustworthy repository or archive will give you with a persistent weblink like a DOI, which directs people to the data. Archive and repositories also give citation formats to make sure data is properly attributed.</p> <p>If you want to use data in future research paper, go ahead. Funders and journals recognise embargo periods for preparing first publications as a legitimate request – but you have to ask, do not assume an embargo is given. Also, embargo periods are usually between six to twelve months long. Anything longer would be based on an exceptional request.</p> <p>If you want to sell access, check your funder agreements, journal policies, and institutional data policy to see if you have the right to sell access. As “employee created works” if there is a commercial use, the School would be interested in asserting its ownership first. Institutional data policies also often insist data be open to the fullest possible extent and stop the assignment of exclusive rights to publishers or external agencies without keeping rights to make the data available for reuse.</p>	
<p><i>Protection</i></p>	<p><i>No way! My data on public attitudes towards the weather is incredibly sensitive and potentially disclosive. I have made promises to destroy my data once the project finishes. Some of what you asked for is confidential. My data have been gathered under complete assurances of confidentiality. We're worried about the Data Protection Act.</i></p>
<p>Requirements to share data are always framed within the law on personal data protection and ethical responsibilities to protect research participants. If you are worried about breaking them, speak to the data library (datalibrary@lse.ac.uk).</p> <p>Funders will not look favourably on a research design that stops data sharing through needlessly restrictive consent agreements or unnecessary promises to destroy data at the end of a project.</p> <p>Make sure you give enough information to participants on data sharing to allow them to make an informed decision on sharing. Remember, data can be anonymised, with access controlled to specific user groups.</p>	
<p><i>Feasibility</i></p>	<p><i>My data is impossible to anonymise. It is impossible to anonymise my transcripts, as too much useful information is lost. I have collected audio-visual data and cannot anonymise them; therefore, I cannot share these data. I am doing quantitative research and this combination of my variables discloses participants' identity.</i></p>

Factor anonymization for reuse into your data management plan, so you can do it early, clearly, and with consistency.

Funders expect you explore possibilities for resolving potential problems with anonymization and sensitivity when it comes to sharing data. If anonymisation is problematic, prohibitively expensive, or would significantly adversely affect the quality of the data, then there are other routes to sharing the data responsibly.

There can be levels of security applied to data that allow reuse to some limited extent.

- licence agreements prevent identification or re-contacting participants.
- secure data services restrict access to data through either virtual online servers or physical locations.

Ownership

*My data collection contains data which I have purchased and it cannot be made public.
There is intellectual property in the data.
We're not sure that we own it.
That data is already published via (external organisation X).
I don't own the data, so can't give you permission.*

Third party Intellectual Property Rights must always be respected – funders, institutions, journals recognise this; and if it is not yours, you cannot give permission.

However, you are expected to make sure you know:

- who owns the data and its terms and conditions of use
- describe how the data can be obtained by others and any limits on access in your research outputs
- or seek permission to share from the IPR holder.



This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

This work is based on elements of:

Louise Corti, Veerle Van den Eynden, Libby Bishop, Matthew Wollard (2014) *Managing and Sharing Research Data: A Guide to Good Practice* (London: Sage) pp.10-11

Christopher Gutteridge, Alexander Dutton, “Concerns about opening up data, and responses which have proved effective”

https://docs.google.com/document/d/1nDtHpnIDTY_G32EMJniXaOGBufjHCck4VC9WGO7jK4/edit#

Laurence Horton (2013) “The five stages to data sharing” Archive and Data Management Training Centre <https://admtic.wordpress.com/category/5-stages-to-data-sharing>