

Exercise: Reasons not to share data

Issues with reuse	
My data is too niche to be of interest/my data has no value	Often, researchers fear that their data might be misused or misinterpreted. This is why proper documentation is so important. If it is clear how the data were created and for what purpose, others will understand the context and know how they can be used. It's impossible to anticipate what uses your data could be put to by others. For example, the weather recordings in old naval logs are being used to monitor climate change now – other communities may see value in your data regardless of how niche it may seem to you.
Others might misuse my data	
Others might misinterpret it and damage my reputation	
My data is impossible to understand without my input	
My data is in the original subject's language, not English	

Issues with publication	
I'm afraid of other researchers using my data to publish ahead of me	As the data creator, you have the right of first use and can apply reasonable embargoes in line with normal practice in your discipline. This gives you time to analyse and publish, but not for an indefinite period. You can still continue to publish on your data while others have access to it, and their publications will also reference your data so you'll get additional credit and citations. It's likely they'll have thought of different uses to you and could even open up opportunities for new research collaborations. Even if you don't produce any digital materials during the course of your research, the outputs you do create will benefit from the publication of a visible, digital record. Large and complex data can be more challenging to share, but data services are available to help. It's worth using these and getting credit for all your research outputs, whether you are required to or not.
I will mine my dataset for a long time to produce several papers	
My dataset is dynamic (e.g. longitudinal dataset)	
My data is too large to share	
My data is not in digital form	
I can't afford to share my data	
I have no data	
My funder does not require me to share data	

Issues with sensitive data	
My research is highly sensitive; I can't make it public	<p>When working with human subjects and investigating sensitive topics, sharing data can be more challenging. It is however still possible! You may need to anonymise or aggregate your data beforehand to protect identities, or if you fear there is still a risk of disclosure, you may opt to use secure data services or data sharing agreements to enforce stricter controls on who can use the data and for what purpose. You may also decide to choose data in certain forms e.g. an anonymised transcript of an interview rather than the associated video recording. Consent agreements should cover options to archive and share data rather than agreeing to destroy data at the end of the project if this is not necessary. Don't assume that participants won't want to participate in open research, many are happy to consent to sharing personal information.</p>
My data contains personally identifiable information	
I combine datasets which, taken together, contain sufficient information to identify participants	
My dataset contains audio-visual data which reveals identity	
My dataset must be destroyed at the end of the project	
Research participants have been given complete assurances of confidentiality	
My data might be the basis of a police prosecution	

Legal issues	
My data is subject to IPR	<p>If you are using secondary data in your research, identify who owns the copyright so that you can ask for permission to publish. Any derived dataset you create may include data belonging to others, and you don't want to be unable to publish your work as you didn't seek full reuse permissions. If legal issues do prevent the publication of your dataset, identify whether parts of the dataset are free of restrictions and consider publishing those. If you discover a commercial application for your research, then you don't want to jeopardise this by publishing. That is not a justification not to publish anything though – it should only affect the data needed to support your application.</p>
I have reused other datasets in my research which I don't have permission to publish	
My data might have a commercial application	