I'm a Sensitive Soul:

How to share sensitive data safely and simply

Sarah Olesen

Australian National Data Service, Canberra, Australia, sarah.olesen@ands.org.au

STATE OF PLAY

Sensitive data has often been excluded or considered an exception to discussions about data sharing. It was believed that sharing sensitive data was unethical or that it is too difficult to do safely. This opinion has changed with more widespread understanding and use of methods to 'de-sensitise' (sometimes 'confidentialise', 'de-identify', or 'anonymise') data. The carrots and sticks for sharing sensitive data have gotten bigger too. As data citation and the use of altmetrics to boost researchers' profiles becomes more commonplace, so too does the incentive to share. Advances in methods of data linkage and harmonisation have also led to a greater demand for data sharing. As have the ongoing needs to better exploit existing data due to competitive and limited funding pools for the disciplines that collect sensitive data – e.g., health and medical research. Finally, major funding bodies and publishers for these disciplines have begun to require or at least 'encourage' data sharing.

But how can this be done? Sensitive, including personal, data is legally and ethically protected, isn't it?....

ANDS GUIDE TO PUBLISHING AND SHARING SENSITIVE DATA

The Australian National Data Service (ANDS) recognised a dearth of resources on sharing for contributors of sensitive data. Legal and ethical obligations surrounding sensitive data sharing were often unclear, in part due to variation across data content, obtained consent, and jurisdiction. In response, ANDS has produced a user-friendly *Guide to Publishing and Sharing Sensitive Data* in consultation with subject-matter experts. Because despite real and perceived barriers to doing so, in most cases **sensitive data can be shared** safely after a certain steps are taken. And there is much to be gained for the researcher, institution, and wider public.

The Guide provides clear descriptions for the terms relevant to sharing sensitive data, including how it is defined in law and research, when 'de-sensitising' is appropriate to allow sharing, data contributors' ethical obligations, data licensing, and sharing via data repositories. The Guide includes an easy-to-follow sequence of instructions as depicted in Figure 1 below (a condensed version from the Guide).

PRESENTATION

This presentation will introduce the ANDS Guide and expand with discussion of the legal, ethical and practical issues faced by researchers and data managers of sensitive data. Case studies will also be provided for illustration.
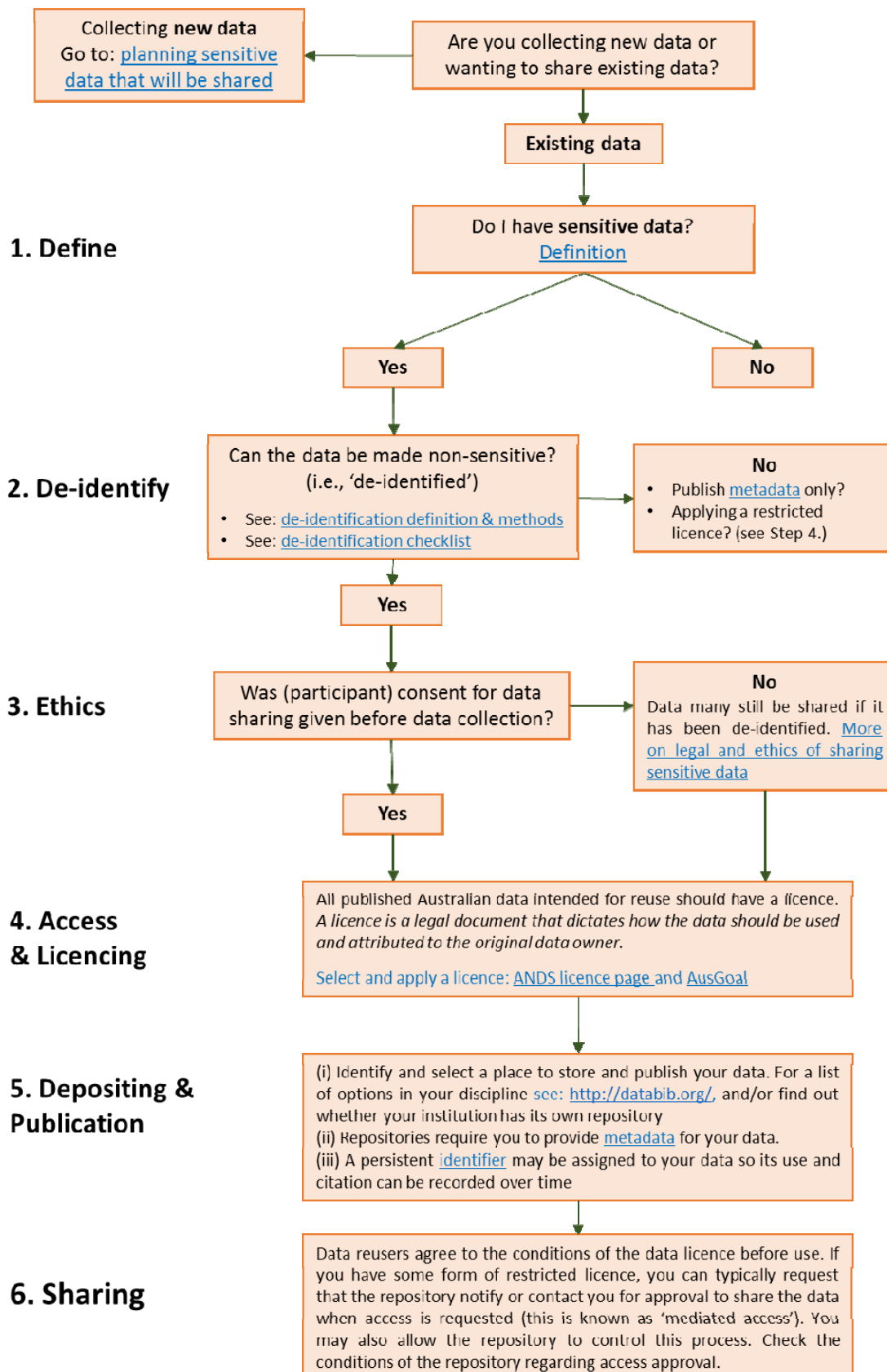
Collecting **new data**
Go to: planning sensitive data that will be shared

Are you collecting new data or wanting to share existing data?

Existing data

1. Define

Do I have **sensitive data**?
Definition

Yes

No

2. De-identify

Can the data be made non-sensitive?
(i.e., 'de-identified')

- See: de-identification definition & methods
- See: de-identification checklist

No
- Publish metadata only?
- Applying a restricted licence? (see Step 4.)

Yes

3. Ethics

Was (participant) consent for data sharing given before data collection?

No
Data many still be shared if it has been de-identified. More on legal and ethics of sharing sensitive data

Yes

4. Access & Licencing

All published Australian data intended for reuse should have a licence. *A licence is a legal document that dictates how the data should be used and attributed to the original data owner.*

Select and apply a licence: ANDS licence page and AusGoal

5. Depositing & Publication

(i) Identify and select a place to store and publish your data. For a list of options in your discipline see: http://databib.org/, and/or find out whether your institution has its own repository
(ii) Repositories require you to provide metadata for your data.
(iii) A persistent identifier may be assigned to your data so its use and citation can be recorded over time

6. Sharing

Data reusers agree to the conditions of the data licence before use. If you have some form of restricted licence, you can typically request that the repository notify or contact you for approval to share the data when access is requested (this is known as 'mediated access'). You may also allow the repository to control this process. Check the conditions of the repository regarding access approval.

Figure 1: Sharing sensitive data steps and sequence

ABOUT THE AUTHOR

Sarah Olesen is Senior Data Management Specialist with the Australian National Data Service and Researcher in the field of mental health epidemiology. She has extensive experience in the reuse of health data.