

## Research Data & Materials Guidelines

# “Making Data Available”

### –guidelines for the sharing, dissemination and publishing of research data

Research Data and Materials Guidelines are provided to complement Research Data and Materials Procedures and practice at VU. (Official procedures are in development.) Guidelines aim to advise researchers on subject specific obligations, standard practices, services and tools.

This document briefly covers reasons for making research data more widely available, if the data is appropriate for sharing, and issues that must be considered when planning to share or publish research data and records. Considerations for sharing data within a research project or collaboration are not covered by this document.

### Why make data available

Data sharing is becoming a key consideration for researchers at the end of projects or as a planned research outcome. There are benefits to providing data – and very good reasons to keep data confidential.

The Australian Code for the Responsible Conduct of Research<sup>1</sup> clearly encourages the communication of research outcomes:

*"Researchers have a responsibility to their colleagues and the wider community to disseminate a full account of their research as broadly as possible" (Section 4.4)*

In addition to the Australian Code, there are a number of benefits in planning to share or publish data.

For example, an increasing number of publishers (including NATURE) require research data to be published and accessible before reviewing articles for submission and peer review.

Publicly-available data is a form of research promotion, as it encourages contacts from research collaborators, from government and from industry. Attaching a DOI (Digital Object Identifier) to your data allows it to be cited, its usage tracked in the same way as journal articles and your research can be more widely recognised. Providing a reference to the published results along with the data encourages further citation of the original published article.

Some data collections can be of significant value to the private sector or industry. In such cases commercial access to data, to search portals, to online analysis, or to data apps might be considered to recover costs. Many successful models include open access to a level of information and commercial access to more detailed data.

Data collections can also be a key resource/facility central to the significance of research groups, attracting recognition, collaborators and further funding.

---

<sup>1</sup> “Australian Code for the Responsible Conduct of Research”, NHMRC, ARC, Universities Australia. 2007.  
<https://www.nhmrc.gov.au/about-us/publications/australian-code-responsible-conduct-research-2018>

There are various models for such “virtual research centres”: including free access to data (relying on increased recognition); collaborator access to data; and access to data on contribution (relying on ever increasing collection significance). All models rely on the public promotion of the data collections, not necessarily public dissemination of data. The first step in promotion is ensuring a description of the collection is discoverable via public registries such as Research Data Australia<sup>2</sup>. Finally, at a global level making data available encourages scientific enquiry, promotes innovation, maximises the transparency and accountability of research, and can reduce the overall cost of research by preventing duplicate data collection. The sharing of research data is considered, by some, the most ethical outcome, maximising the research benefit for the impact on human and animal participants. But for human data in particular this must take into consideration privacy, informed consent and the risks of inappropriate use.

### **Ethics, Privacy, Sensitive Data and Integrity**

The NHMRC and ARC strongly support the dissemination of data that results from government funding. However, their means to achieve this end is based upon open access to *publications*. This is particularly significant, because publication, especially those in a refereed journal, carries with it safeguards relating to the use of data. In general these map closely onto the ethics applications and approval that has been obtained, and in many cases journal editors now ask to see proof of this.

There is also a push to share unpublished, detailed, anonymised or source data. This must be approached with great caution for research involving or impacting on humans, or for data including private, confidential, sensitive or potentially harmful information. To take an example the Australian National Data Service (ANDS) provide a guide on “Ethics, Consent and Sharing Data”<sup>3</sup>: *Research data – even sensitive and confidential data – can be shared ethically and legally if researchers pay attention, from the beginning of research, to four important aspects:*

- *including provision for data sharing when gaining informed consent*
- *protecting people's identities by anonymising data where needed*
- *considering controlling access to data*
- *applying an appropriate licence*

ANDS also advise<sup>4</sup>:

*Encourage data sharing by ensuring that ethics forms include relevant questions ...*

*Provide advice to researchers about designing their research so that the data can later be shared ...*

*Anonymise the data ...*

*Add a question to your ethics form: "Is there any reason NOT to share the data from this project?"*

This advice suggests an “opt out” system where it becomes usual to share data, rather than not to share it.

While it is certainly preferable to resolve issues relating to informed consent at the beginning of a study when the ethics process is being addressed, several matters need attention.

---

<sup>2</sup> <http://researchdata.ands.org.au/>

<sup>3</sup> <http://ands.org.au/guides/ethics-working-level.html>

<sup>4</sup> <https://www.ands.org.au/working-with-data/sensitive-data/ethics-and-data-sharing>

1. At the present time, is it possible to provide clear detail as to the purpose of the current and future research, and the time frame in which the research will take place? If not, blanket consent becomes meaningless. The reasons for this are twofold:

- a) The potential risks and ethical issues inherent in the research are not able to be assessed. Nor is it possible to assess who will see the data in its raw state, even if it has been made anonymous.
- b) It is not possible to give informed consent without this information.

Currently there is much emphasis (rightly) on informed consent as the protector of privacy and confidentiality. What then follows is the assumption that if these are protected then data can be shared. There are, however, other issues that relate to informed consent. The most significant of these relates to outcome of research. Namely, participants have the right to understand the use to which their data is to be put, and to feel comfortable with the potential outcomes. It is not really possible for this to happen if there is no knowledge of what these outcomes might be.

While an important aspect of informed consent is the protection of privacy and confidentiality, the most significant aspect relates to use and research outcomes. Participants have the right to understand the how their contributions will be used, and to feel comfortable with the potential outcomes. This is not possible if there is no knowledge of what these outcomes might be.

If in doubt, please contact the Office for Research to seek advice.

### **Protecting your Intellectual Property**

For the purposes of this document the word data is used in its widest sense to mean information that is captured, analysed and reported on from a variety of research activities. It may be the raw results of an experiment, simulation or survey or it may be extracted from existing models, documents or databases. Before sharing any data that you have access to or have created, there are a few things you should think about.

#### **Is the data yours?**

Any data you make available to others may be subject to copyright, to privacy and confidentiality agreements and could be commercially valuable IP for further development – either as the basis for new research or for commercial development activities. The first thing you must be sure of is that *you have the right to make the data available – that is, make sure you own or have the rights to the data.*

Remember this does not necessarily mean that because you created the data that you have the right to communicate it – for example, making available explicit, personal information regarding human subjects that may have been involved in your research is poor ethical practice. Similarly, contract research activities where the terms of the project cedes ownership of all project IP to the customer may also preclude you from sharing.

You can check with the VU library, the Office of Innovation and Intellectual Property or with VU Legal Services if you're not sure.

#### **The data is mine – but what might I want to do with it?**

Having established that you have the right to disseminate the data, your next step is to think about what you might want the data for in the future. If the data has commercial potential, you may wish to retain confidentiality regarding some or all of the information – which means you may wish to

modify the form of your material before publishing. Again, check with the Office of Innovation and Intellectual Property and/or with VU Legal Services if you're not sure.

### Releasing Data

Having no ethical, legal or commercial development reasons to prevent dissemination, you may be ready to release your data. But there are still a few things to consider. For example, it would be wise to circulate the most recent and complete form, rather than a draft, or incomplete data set. You may wish to consider what form the data is in (is it digital, written, fabricated?) and what is likely to be the most common access method (for example, file size and type if it is digital information). Similarly, you may need to think about how you will categorise your data and how you will make it available. You may wish a royalty or licence fee for those seeking to use your data – so some thought about what it is worth is also required. The College Research Facilitation Managers can help here and direct you to further resources and individuals that you may require.

Lastly, you need to provide some documentation together with your data as it's difficult to know where it may end up once it has left your control. This should include the following:

- A statement of ownership, authors, contributors or other attributions and acknowledgements.
- A license describing any restrictions on use or clarifying permission for reuse. This should include if others must acknowledge VU and contributors in subsequent outcomes. Consider including instruction on "how to cite" the data, for example the publication in which the data was first published.
- A disclaimer is highly recommended.
- A description of the data itself. As a minimum include a brief description of the contents, context, and any known limitations, issues, uncertainties and accuracy.

See "[Appendix B: Example Ownership and License Statement, and Disclaimer](#)" for an example as well as a recommended disclaimer.

### Where to Release or Share Data

Consider depositing your data in a discipline repository or well established archive. This will facilitate data discovery, preservation, access and distribution – all generally managed by the repository/archive curator or self service in the case of open access repositories. This can save your time in management of data and facilitating requests. Data can be shared informally via emailing or posting to social websites, but such methods are not long-term solutions.

There are several sites that can help you discover national and international repositories and archives. Use well known repositories for your discipline if possible.

- <http://databib.org/>
- <http://www.re3data.org/>
- <http://www.ada.edu.au/> - Australian Data Archive (Social Science, Historical, Indigenous, Longitudinal surveys of the Australian population, Qualitative, Crime & Justice)

Alternative mechanisms for publishing digital data are:

- <http://FigShare.com/>
- VU's Institutional Repository (contact the Library)
- Self-dissemination through dedicated team or project site/portal. (contact Office for Research, eResearch regarding national cloud resources)

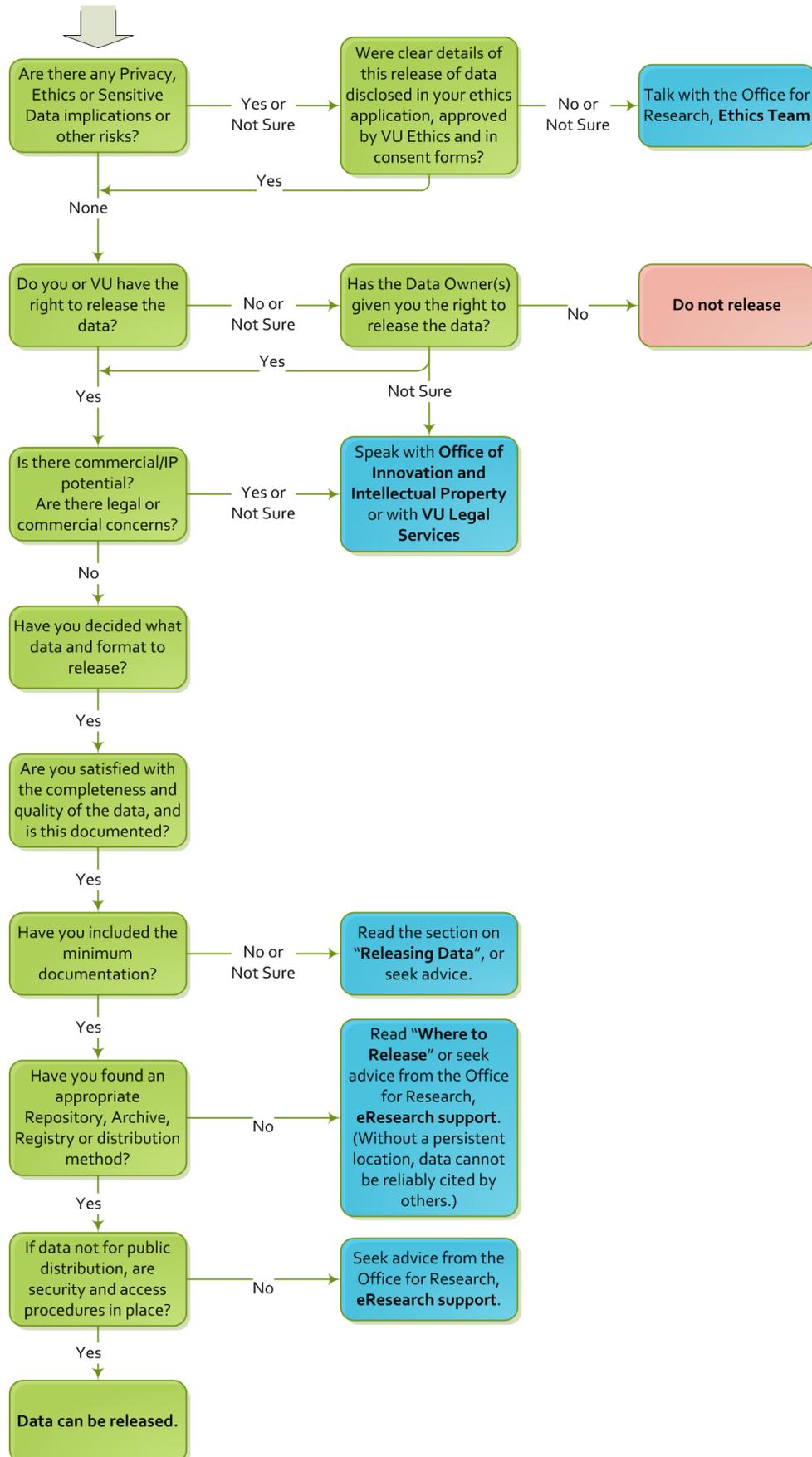
Research Data Australia is the national registry that enables research datasets, collections or portals to be promoted and discovered. This can be used by any VU researcher to promote datasets or collections, physical or digital, shared via the above mechanisms or informally. Research Data Australia does not retain data, only descriptions of datasets and how others can gain access. For more information see the VU service page <https://www.vu.edu.au/researchers/research-contacts-support/research-data-reporting-eresearch> or visit Research Data Australia to discover existing VU collections and more <http://researchdata.andcs.org.au/victoria-university/> .

## **ACKNOWLEDGEMENT**

Victoria University has drawn on the following sources of information in drafting this procedure:

- Professor Warren Payne (PVC-R), sponsor
- Lyle Winton and Jo Dalvean, authors and editors
- John Rabling, author of Intellectual Property section
- David Kay, IP and commercial advice
- Deb Zion, author of Ethics and Integrity section
- ANDS and CSIRO for their leadership and inspiring sections of this document.

## APPENDIX A: Checklist for making research data available



**Appendix B: Example Ownership and License Statement, and Disclaimer** The following is an example of an actual VU dataset ownership and license statement that has been made available, a requirement for publishing in Nature. In this case the license is open access (Creative Commons) requiring only that the work be acknowledged via a citation. This statement was approved by the Office for Knowledge Exchange in 2012. The license was obtained from AusGOAL, the Australian Government Open Access and Licensing Framework. The following is a disclaimer and legal notice in use on the data.vu.edu.au site.

Copyright (c) 2013, Roger Bodman, Victoria University - all rights reserved.

This work is licensed under a Creative Commons Attribution 3.0 Australia License.

[http://creativecommons.org/licenses/by/3.0/au/deed.en\\_US](http://creativecommons.org/licenses/by/3.0/au/deed.en_US)

In sharing or adapting this work you must acknowledge this work.

How to cite this work:

Bodman, R. W. et al. (2013), "Reduced uncertainty in temperature projections using carbon cycle and climate observations" Nature Climate Change, Vol 3, p725-729. doi:10.1038/nclimate1903.

### **Disclaimer and Legal Notice**

All rights in materials (files, data, data sets, databases and other content) made available belong to Victoria University unless otherwise specified. All materials are made available for use on a non-exclusive, non commercial basis, unless otherwise specified. Full terms of use may be set out in the README or LICENSE files/statements available within each Data Set. Please contact the nominated contact person for each individual Data Set.

In using materials you do so at your own risk and you acknowledge that:

- The information and material made available are considered research data and are provided “as is”;
- Materials may contain information and statements based on research and collected by Victoria University and third parties;
- Materials may contain information and statements that are not professional, scientific, medical, technical or expert advice and that are subject to the usual uncertainties of advanced research;
- Users bear all responsibility and liability for their use of data, and for any loss of business or profits, or for any indirect, incidental or consequential damages arising out of any use of, or inability to use, the data, even if Victoria University were previously advised of the possibility of such damages, or for any other claim by you or any other person;
- Victoria University makes no representations or warranties of any kind, express or implied, including implied warranties of fitness for a particular purpose or merchantability, or with respect to the accuracy of or the absence or the presence or defects or errors in data, databases of other information. In particular, materials may contain raw data that is uncorrected; and
- The information and materials do not imply the expression of any opinion whatsoever on the part of Victoria University.