DEPOSITING DATA WITH THE AUSTRALIAN DATA ARCHIVE

User Guide No.3

Introduction

The Australian Data Archive (ADA) is a consortium of leading national Australian universities, managed by the Australian National University (ANU) and with nodes at the University of Melbourne, University of Queensland, University of Technology Sydney and University of Western Australia. ADA provides a national service for the collection and preservation of computer readable data relating to social, political and economic affairs and to make these data available for further analysis. The core of this service is the Australian Data Archive website.

An overview of the Archive

The Archive has been arranged to enable fast access for users. Our catalogue of holdings is accessible as a single archive (ADA), with a series of sub-archives reflecting the nature of the data and/or major thematic divisions. These are:

- Social Science – predominantly survey or polling based quantitative social science data

- Historical – an archive of Australian census data tables from 1834 to the present day

- Indigenous – A thematic archive bringing together research data about Aboriginal and Torres Strait Islanders

- Longitudinal –major longitudinal cohort and panel surveys of the Australian population

- Quallitative – a new collection which provides specialist data archiving and access services to qualitative researchers

- Crime & Justice – major collections of data in crime, law and justice, including criminal justice administrative data

- International – a central point of access for links to international data sources around the world

This organisation means that you can browse and search the full ADA collection, or browse and search each subarchive individually.

Depositing Data

The Australian Data Archive urges all researchers and research groups to take the opportunity of having their data professionally archived. There are two main reasons for depositing your research data with the Australian Data Archive: to ensure that your data is preserved for the future, and to enable it to be shared with others for secondary analysis.

The ADA data deposit pages provide advice to researchers on:

- Why you should deposit your data: the benefits of data archiving and sharing

- How to deposit your data with ADA

- Determining conditions for access to your data

We have provided a Deposit Form and Licence Form (at the end of this guide) to assist us in archiving your data. These forms will also assist you in getting your deposit together.

Steps to follow to deposit your data with ADA include

1. Collecting your data files

2. Collecting your documentation

3. Filling out the ADA license and deposit form

4. Depositing your data with ADA

Why deposit your data?

There are three broad areas of benefit to depositing your research data with the Australian Data Archive: benefits for you as a researcher, for the research community, and for the public at large. Archiving your data helps to ensure that your data is preserved for the future, and enables it to be shared with others for secondary analysis.

Benefits to you

There are significant benefits to you as a researcher in depositing your data. Advantages include:

- Your data will be stored indefinitely in a safe place, where it is easily accessible, and in a form that can be understood long after the study was originally conducted. This is useful if you are obliged to store the data for several years following the publication of results.

- The potential impact of your data is greatly increased. Other researchers, even from other disciplines, will be able to use your data (subject to the conditions you specify) to investigate issues that you may not have the desire or resources to pursue.

- Anyone who publishes results from your data is expected to cite the data file in their publications. This will result in increased citations for you and your colleages.

- Distribution of the data and documentation is managed without expense to the person who originally conducted the research. If other people become interested in the data, you will not need to be concerned with the cost and inconvenience of providing a copy for them.

- Publicity for your data: your data is published in the ADA data catalogue, raising the profile of your research.

- Sharing data through ADA allows you to get in touch with other researchers with similar research data interests to you.

- You can find your own data in future. Should you have difficulties in locating or accessing your data in the future, you will have a preserved version of the data held with ADA.

The NHMRC/ARC Australian Code for the Responsible Conduct of Research (ACRCR) also strongly promotes data preservation and sharing. Section 2 of the ACRCR relates to the management of research data. It suggests:

- both institutions and individual researchers have an obligation to preserve research data, in a secure environment and allow others to access it where appropriate,

- datasets should be kept by departments, institutional repositories or national repositories, rather than only by the researcher, and

- the value of material for other research where the research is difficult or impossible to repeat must be considered.

Benefits to the Research Community

Sharing data builds a stronger research community by:

- Providing researchers and students access to high quality data.

- Allowing verification, replication and extension,

- Supporting multidisciplinary, multi-institution research,

- Reducing the duplication of similar research

It is important to the scientific method that social science data generated from research projects be available for replication, verification and extension. A researcher can access data from ADA, verify the results of the original investigator, conduct analyses not performed by the original investigator, or use the data, survey instruments and methods to guide their own original research.

The costs of data collection are high, and beyond the reach of most students. For those with limited access to research funding. Archived data allows the students to learn about research method and

analysis in a practical way.

The importance of sharing has als been recognised by funding bodies in Australia and internationally (including UNESCO, the UK Economic and Social Research Council, the US National Science Foundation, NHMRC and the ARC) and is now considered a vital part of the research process.

Benefits to the Public

The dissemination of research data and results also has potential benefits for the general public. For example, the Australian National University (in their Position Statement on Open Access) has identified that:

> "The University has an obligation to share research outcomes with the global community, many members of which do not enjoy the same level of investment in research experienced in Australia. The ability to participate and share in scientific and cultural advancement is a declared human right.

> Traditional publication channels do not necessarily place scholarly works into the public domain promptly and at low cost to the user. The University supports prompt open dissemination where possible to avoid these barriers."

Publishing data through the Australian Data Archive enables the public to gain access to research data that would otherwise be inaccessible.

How To Deposit Data

There are 4 simple steps to depositing your data with ADA.

1. Collect together all the data files

In the course of your normal research activities, ADA archivists often find that you have already prepared your data in a form that would be suitable for depositing with ADA.

In preparing your data files for deposit, you may want to think about:

- confidentiality - removing variables like peoples' names, addresses and phone numbers. (We will let you know if we find something that we think is a risk)

- naming conventions - have you labelled your data files and contents (such as variables) in a way that makes them readily understandable to others?

- created and derived variables - if they are important to replication of your work, can you provide the syntax that created it? If you can't describe how you created them, is it still important to archive them?

- data formats - are your data in an open or commonly used format? ADA does have a set of preferred data formats for deposit, but we can handle most data formats. (See Appendix One for an overview of suggested formats). ADA can provide some data format conversion, but this is dependent on the availability of ADA archivists and appropriate software and hardware.

If you're not sure about whether your data is in an appropriate format, please contact us.

2. Collect together all the relevant documentation

You may want to include:

- a questionnaire and any other materials used like show cards

- technical reports

- any publications you want referenced in the documentation

- other materials that may be helpful to a researcher in analysing and interpreting your data

It is preferable if you can provide these materials in electronic format where possible.

3. Complete the ADA deposit and license forms

To deposit data with the Australian Data Archive, please complete the following forms:

- ADA Deposit License Form: http://ada.edu.au/documents/ada-data-deposit-license-form

- ADA Study Description Form: http://ada.edu.au/documents/ada-deposit-study-description-form

These forms assist us in archiving your data (see links at Appendix Three) and making it available to others for reuse. These forms will also assist you in getting your deposit together.

It is essential to complete the ADA Deposit Licence Form as it gives us your permission to archive the data and distribute to others. This includes determining the access conditions for your data (see further below for information on Setting Access Conditions).

The Deposit form provides us with information on your research so that we can provide an adequate description of the data, the methods use, documentation and other information useful to other users of the data. We will contact you if we find there is further information that we need to create complete documentation for your data.

4. Deposit the data, documentation and forms with ADA

There are a number of ways to transfer your deposit:

- directly to an ADA data archivist

- online - ADA has developed specific methods for online data transfer, including FTP and a secure web deposit service.

- by email - suitable for smaller data files

- by post - your electronic files can be copied onto a CD-Rom, USB drive or similar, and mailed to us. It's a good idea to let us know that it is coming so we can let you know when it is in our care. It is good practice to password-protect / encrypt data files that are sent by post

To determine the best method for the deposit of your data, please contact us at ada@anu.edu.au.

Setting Access Conditions

The purpose of depositing data in the Australian Data Archive is to enable the data to be made available to other researchers who may wish to analyse them. Under ADA's deposit protocols, the conditions under which the data may be made available by the Archive to other researchers are determined by you as the Principal Investigator depositing the data. Catalogue information about a study and it's contents (known as metadata) is freely accessible to enable researchers to locate relevant data (and thus not subject to conditions).

The procedures by which a user may access data are set out in our Access Conditions (see Appendix Two). In order to access any data for online analysis or for download a user must first register with ADA. The particular form of access conditions is set three possible levels of access: General, Restricted or Special Access.

On the ADA license form you can choose one of the two standard access conditions:

- General access: There are no additional restrictions on access to the data or publication of results, and the depositor does not wish to be informed of the use being made of the data.

- Restricted access: The depositor, or an authorised representative, wishes to be informed by the Archive of each request to use the data in order to give or withhold permission.

In addition, you can liaise with ADA to determine special access conditions. For example, a user interested in accessing data from a study may be required to obtain your permission, or that of an authorised representative (such as the ADA National Manager), in writing before publishing any interpretation of such materials.

As a depositor, you can also contact the Archive at any time to find out who is using your data, in order to comment on that use and make contact with colleagues of similar interests. Users are notified that their information may be shared with depositors at the time they request data from a study.

Embargoes

Researchers are often interested in publishing from their data prior to releasing the data for access to others. As a depositor, you can request that an embargo period be imposed on your data, whereby no access to the data would be permitted until after the date you specify. At the end of the embargo period, the data may be released under the access conditions for the study.

As a depositor, you can also set embargo or access conditions to suit their particular concerns about how the data should be used in the future. If you would like to ensure that other researchers do not pre-empt your findings, you can stipulate that the data be made available only after you have had time to publish your results.

Contact us:

If you require further information on depositing your data, would like additional advice, or have further questions, please don't hesitate to contact the archive.

To contact ADA, please see our details below. The preferred form of contact for ADA is via email.

Email:

ada@anu.edu.au

Telephone:

(+61) 2 6125 2200

Facsimilie:

(+61) 2 6125 0627

Street and Postal Address:

The Australian Data Archive

The Australian National University

18 Balmain Lane

Acton ACT 0200

Appendix One: Preferred Data Deposit Formats

Quantitative data

Quantitative data is preferred to be deposited in one of the following formats.

Quantitative Data	Preferred Format	Other acceptable formats
Statistical packages	SPSS, Stata	SAS, R
Spreadsheet data	CSV	Excel
ASCII text	CSV, Tab-delimited text, fixed format text	

Other formats, such as database formats (eg. Microsoft Access, MySQL, PostGres) should be discussed with ADA staff prior to deposit.

Qualitative data

For qualitative data, possible data formats vary significantly. For this reason, ADA has developed specific preferred formats, partly based on those identified by the UK Data Archive (reproduced with permission).

Qualitative Data	Preferred Format	Other acceptable formats
Textual	eXtensible Markup Language (XML) marked-up text according to an appropriate Document Type Definition (DTD) or schema; Rich Text Format (.rtf); Plain text data, ASCII (.txt)	Hypertext Markup Language (HTML); Widely-used proprietary formats e.g. Microsoft Word (.doc/.docx); Proprietary/software-specific formats such as NUD*IST, NVivo and ATLAS.ti
Digital image data	TIFF (uncompressed)	JPEG (.jpeg, .jpg); Adobe Portable Document Format (PDF/A or PDF); Raw image format (.RAW); Software-specific formats (such as, for example, Photoshop .psd files) may be acceptable, but contributors should contact ADA for advice before file upload
Digital audio data	Free Lossless Audio Codec (FLAC)(.flac); WAV file (.wav)	MPEG-1 Audio Layer 3 (.mp3); Audio Interchange File Format (AIFF)(.aif)

Digital video data	JPEG 2000; MPEG4	Contributors should contact ADA for advice before file upload.
Documentation	RTF (.rtf); PDF/A or PDF; HTML (.htm); Open Document Text (.odt)	Plain text (.txt) Widely-used proprietary formats e.g. Microsoft Word (.doc/.docx) or Excel (.xls/.xlsx), are acceptable but offer less long-term security; XML marked-up text according to an appropriate DTD or schema, e.g. XHMTL 1.0

Appendix Two: Conditions of Access to Data from the Australian Data Archive

(Source: http://www.ada.edu.au/ada/access-conditions)

The Australian Data Archive is committed to providing open access to Australian and international research data for research and education purposes. This open access commitment however is balanced against our obligations to the original participants in these research studies. Data provided to the Australian Data Archive has been collected from research participants following research ethics requirements on the depositor who produced the data. These requirement place obligations on the researchers and ADA for appropriate use of the data for secondary purposes.

Access Categories

Data from the Australian Data Archive is made available under three broad categories of access. The choice of access category is made by the depositor of the data:

- GENERAL ACCESS: studies with no restrictions placed on data access

- RESTRICTED ACCESS: studies where the depositor, or an authorised representative, wishes to be informed by the Archive of each request to use the data in order to give or withhold permission.

- SPECIAL ACCESS: studies where the depositor has included additional special access conditions. For example, the user may be required to obtain the permission in writing of the original depositor of the data, or an authorised representative, before publishing any interpretation of such materials.

Access to data at each of these levels has an associated user undertaking that you must complete in order to be granted access.

Access Levels

The Archive then provides access to data in the Australian Data Archive at three levels, depending on the access category of the study:

- BROWSE: view metadata and summary statistics for data files

- ANALYSE: use the ADA Nesstar analysis system to run cross-tabulations, correlations and linear regressions

- DOWNLOAD: export the raw data files to your own computer

Your access to data at any point is then a combination of the access category of the study and the level of access to the data that you are seeking. Each of these access levels incorporates additional expectations on you regarding secure and appropriate use of the data. These can be summarised as follows:

	Access Category		
Access Level	General	Restricted	Special
Browse	Open	Open	Open
Analyse	Open to registered users	Application required (Depositor approval)	Study-specific
Download	Application required (ADA approval)	Application required (Depositor approval)	Study-specific

The process for accessing ADA data then depends upon the level of access (browse, analyse or download) that you require. Browsing access is available for any study, while access to analyse data under a general access license is open to registered ADA users - click on the "Register" button at the top right of this page to create your account and log in to the ADA website.

Applying for access to data

To apply for access to general or restricted access datasets, you need to complete a Data Access Request.

If you know the name or ADA ID number of the study you are interested in, you can make a direct request through your "My Account" tab - you must be logged in in order to make a request. Alternatively, you can follow the Analyse/Download links from an ADA study, which will provide an automated request form for you to complete if you have not been previously approved.

Costs of Access

The costs of access to data from the Australian Data Archive are determined on the basis of the institution you belong to (eg. as a staff member, student, adjunct, etc.)

1. Staff and Students of ACSPRI member institutions

The Australian Consortium for Social and Political Research Inc. (ACSPRI) has negotiated free access to ADA data for staff and students of all its member organisations. Your organisation's ACSPRI membership is determined automatically when you enter the name of your institution in your user registration. Most Australian universites and a number of Government agencies and non-government organisations have ACSPRI membership (see the online list of ACSPRI Member Institutions). For information on joining ACSPRI see their website information on Becoming a Member

2. Staff and Students of ADA international partners

ADA is a member of the International Federation of Data Organizations for the Social Sciences (IFDO) and holds data exchange agreements with a number of different data archives and services around the globe. This includes staff and students from member institutions of the Inter-University Consortium for Political and Social Research (ICPSR) and the Council for European Social Science Data Archives (CESSDA) member institutions. You can contact us or your local data service to find out whether your affiliation entitles you to free access to ADA data.

3. All others

You are required to apply for each dataset individually and pay a $1000 + GST (Aust.) administrative fee per dataset.