

Exercise 1: Defining a collection

Two examples of collections with subcollections and associated materials are shown below. Decide what level of collection should be described in Research Data Australia for each one..

Census data from the Australian Bureau of Statistics

<http://www.abs.gov.au/websitedbs/d3310114.nsf/home/census+data>

Census Data

Access Census Data Online  [Quality Declaration](#)

This page provides you with access to all Census data that is available online. Census data is displayed based upon your selected location and topic of interest, and can be viewed in different formats including thematic maps or tables. You can search for data based primarily on your selected location or topic, or you can go straight to one of the online tools to access data in the format you need:

Find Census data by Location:

[2006 Census](#) - [2001 Census](#) - [1996 Census](#)

Find Census data by Topic:

[2006 Census](#) - [2001 Census](#)

Find Census data using our Online Tools:



QuickStats - provide a summary of key Census data relating to persons, families and dwellings. QuickStats cover general topics about your chosen location and include Australian data to allow comparison.

[2006 QuickStats](#) - [2001 QuickStats](#)



MapStats - provide thematically mapped Census statistics for your chosen location. The maps illustrate the distribution of selected population, ethnicity, education, family, income, labour force, and dwelling characteristics.

[2006 MapStats](#) - [2001 MapStats](#)



Census Tables - are designed for clients who are interested in data on a specific topic. Census Tables provide individual tables of Census data for a chosen location in excel format.

[2006 Census Tables](#) - [2001 Census Tables](#)



Community Profiles - are available as six distinct profiles of key Census characteristics relating to persons, families and dwellings, and covering most topics on the Census form for your chosen location.

[2006 Community Profiles](#) - [2001 Community Profiles](#)



CDATA Online allows you to create your own tables of Census data on a range of topics such as, age, education, housing, income, transport, religion, ethnicity, occupation and more. All Census geographies will be available, from a single collection district to an entire state/territory or Australia. [More info...](#)

[2006 CDATA Online](#)



TableBuilder - Census TableBuilder is an online tool which allows you to create your own tables of Census data by accessing all variables contained in the Census Output Record File for all ABS geographic areas. It is a charged subscription service.

[More info...](#)

[Access TableBuilder](#)

What collections should be shown in Research Data Australia? (no right answer)

Extracts from the *Longitudinal Study of Australian Children*

<http://www.aifs.gov.au/growingup/data/index.html>

File structure

For the Wave 3 data release, the following datasets have been made available:

- Six datasets comprising the main datasets for each wave and cohort (lsacgrb04, lsacgrb2, lsacgrb4, lsacgrk4, lsacgrk6 and lsacgrk8)
- 12 Time Use Diary datasets:
 - One cleaned datafile with problematic cases deleted for each Wave and cohort (diaryb0, diaryb2, etc.)
 - One with the cases deleted from the above files after cleaning for each wave and cohort (poortudsb0, poortudsb2, etc.)
 - One with all cases and no data cleaning performed on them for each wave and cohort (ucdiaryb0, ucdiaryb2, etc.)
- Three Medicare Australia Datasets representing information from the 3 Medicare Australia databases the information was drawn from (mbs, pbs and acir)
- Two household composition datasets, 1 for each cohort (hhgrb, hhgrk)
- Two Wave 2.5 datasets, 1 for each cohort (lsacgrb3, lsacgrk7)

This structure has been used to reduce the size of the main datasets and because some data are better formatted using more than one record for each child.

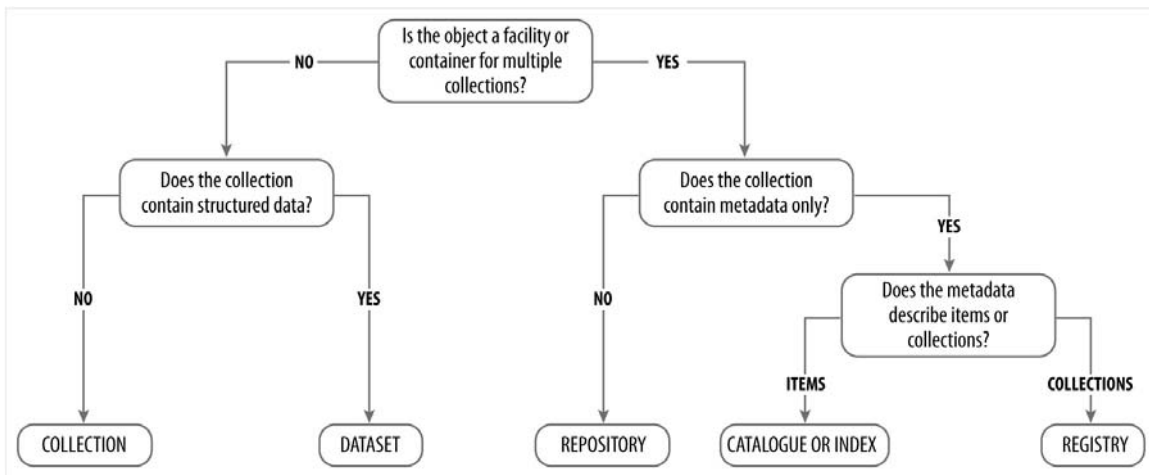
How data files are provided

All data are provided in both SAS and SPSS formats, however users can transfer the data to other formats if they wish. The CD-ROM and/or website also includes extensive data documentation, including this document, marked-up questionnaires and variable frequencies. The data files and the other documentation are discussed in detail in later sections of this document.

What collections should be shown in Research Data Australia? (no right answer)

Exercise 2: Making decisions about collection types

Use the following decision tree to assign a collection type (collection, dataset, repository, catalogue or index, registry) to the example collections. The information provided may not be enough for a definitive answer.



Agincourt Reef Automatic Weather Station (Great Barrier Reef)

Brief description: This dataset contains meteorological data from Agincourt Reef, Great Barrier Reef, from November 1989 with breaks. The weather station is located on Quicksilver's northern pontoon. [record from Australian Institute of Marine Science]

Collection type? _____

Australian Archive of Jewish Music Collection

Brief description: This collection of audio files available through the Monash University ARROW Repository is a subset of the unique Australian Archive of Jewish Music Collection located in the School of Music in the Performing Arts Building of Monash University. It covers pre-1950 recordings of Hebrew and Yiddish songs and performance representing the Jewish music culture of Australia and Asia. [record in Monash ARROW repository]

Collection type? _____

Australian Research Online (ARO) OAI-PMH repository

Brief description: The Australian Research Online (ARO) OAI-PMH repository provides access to records harvested from Australian university and government research repositories in addition to several other collections of Australian research.

Collection type?

X-ray crystal structure of the fibrinolysis inhibitor {alpha}2-antiplasmin

About: The serpin alpha(2)-antiplasmin (SERPINF2) is the principal inhibitor of plasmin and inhibits fibrinolysis. Accordingly, alpha(2)-antiplasmin deficiency in humans results in uncontrolled fibrinolysis and a bleeding disorder. alpha(2)-antiplasmin is an unusual serpin, in that it contains extensive N- and C-terminal sequences flanking the serpin domain. The N-terminal sequence is crosslinked to fibrin by factor XIIIa, whereas the C-terminal region mediates the initial interaction with plasmin. To understand how this may happen, we have determined the 2.65A X-ray crystal structure of an N-terminal truncated murine alpha(2)-antiplasmin. [record from Tardis Federated Diffraction Image Publication Repository ,contains data and metadata]

Collection type?

Tardis Federated Diffraction Image Publication Repository

About: TARDIS is a multi-institutional collaborative venture that aims to facilitate the archiving and sharing of raw X-ray diffraction images (collectively known as a 'dataset') from the Australian protein crystallography community. [contains data and metadata]

Collection type?

Exercise 3: Identifying data to share with Research Data Australia

In the following example, identify information to share with the ANDS Collections Registry to support discovery through Research Data Australia by *ticking* items to be included.

The essential criterion is whether or not the information will help, firstly, discovery and secondly, decisions about re-use of the data.

Example from Australian Social Science Data Archive

[http://search.assda.edu.au/search/index.jsp?queryString=queryAnd:\[au.edu.anu.assda.ddi.01076\]+queryFieldType:\[both\]+queryField:\[Dataset+ID\]+server:\[ASSDA\]](http://search.assda.edu.au/search/index.jsp?queryString=queryAnd:[au.edu.anu.assda.ddi.01076]+queryFieldType:[both]+queryField:[Dataset+ID]+server:[ASSDA])

Dataset: Families, Social Capital and Citizenship Survey, 2001

Dataset ID: au.edu.anu.assda.ddi.01076

Dataset name: Families, Social Capital and Citizenship Survey, 2001

Abstract

The Families, Social Capital and Citizenship project aims to examine levels of social capital associated with varying family circumstances, assess the importance of social capital in shaping patterns of family engagement with the economy, polity and community, and assess the relative importance of different elements of social capital (including trust, reciprocity and the size and structure of networks) to these outcomes.

Variables include: institutional ties; involvement with different community groups; the extensiveness of informal ties with family, friends and neighbours, levels of trust and reciprocity within these networks, the receipt or provision of emotional, practical and financial support within these networks, the use or provision of care for the elderly, the disabled or children; trust and confidence in different institutions; and feelings about ones neighbourhood.

Background variables include: age, gender and relationships of persons in the household; respondent country of birth and languages spoken; occupation, education, income and income sources; parent's education; and partners and friends education and occupation.

Topic categories: Australian Studies; Sociology, Culture; Social attitudes, values and behaviour

Keywords: Community involvement; Education; Employment; Ethnic Groups; Groups; Income; Family; Volunteer workers

Principal investigators: Jody Hughes (Australian Institute of Family Studies); Wendy Stone (Australian Institute of Family Studies)

Last data collection date: 23 April 2001

Country: Australia

Universe: Australians aged 18 years and over with a telephone number listed in the electronic White Pages.

Time method: One-time cross-sectional study

GUMBOOTS FOR THE DATA DELUGE, SYDNEY, 21 APRIL 2010

Sampling procedure: One-stage stratified or systematic random sample

Mode of data collection: Telephone interview

Special permissions: The depositor has requested that the user apply for permission to access these data. Please complete the Application to Access Restricted Data Form (<http://assda-nesstar.anu.edu.au/webview/ASSDADData/forms/ASSDA.FORM.ACCESSB.pdf>) and return it to ASSDA who will contact the depositor on your behalf.

Related Material:

[Online Codebook](#)

[Codebook](#)

[Questionnaire](#)

[Report](#)

Exercise 4: Creating relations in Research Data Australia

In this scenario, Professor Smith was funded by the Australian Research Council to undertake a research project called “Urban Renewal and Drought”. Outputs of the project were a publication and a dataset containing data about open space and block size in housing developments.

The following objects are described and linked together in the ANDS Collections Registry;

- Professor Smith (party / person) participates in the project activity, and collects and manages the dataset
- Australian Research Council (party /group) funds the activity
- Urban Renewal and Drought Project (activity) has the dataset as an output (also has publications as outputs)
- Housing development open space and block size 2008 dataset (collection) supports a service (relation already shown)
- Urban renewal data search and supply service (service)

Draw and label the appropriate links between each related object's KEY in the following diagram:

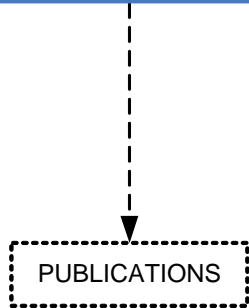
- funds/isFundedBy
- isParticipantIn/hasParticipant
- isManagedBy/isManagerOf
- hasCollector/isCollectorOf
- hasOutput/isOutputOf

KEY: http://nla.gov.au/nla.party-536838
PARTY (GROUP): Australian Research Council

KEY: http://nla.gov.au/nla.party-xxxxxx
PARTY (PERSON): Professor Smith

KEY: http://local.unique.key
ACTIVITY (PROJECT): Urban Renewal and Drought Project

KEY: hdl:102.100.100/xxxxxx
COLLECTION (DATASET): Housing development open space and block size 2008 dataset



Supports/isSupportedBy

KEY: http://local.unique.key
SERVICE: Urban renewal data search and supply service

Note that links to publications are links to objects outside the registry, and are made from the Related Object element, not using the internal registry relations.