



HUMANITIES, ARTS & CULTURE DATA SUMMIT

Steve McEachern
Australian Data Archive, ANU



14-15 March 2018, Canberra
humanities.org.au | #HACDS2018



The Australian Data Archive as infrastructure for the HASS community

Dr. Steven McEachern

Director, Australian Data Archive

ANU Centre for Social Research and Methods

ADA in Brief

- The Social Science Data Archive (now ADA) was set up in 1981, housed in the Research School of Social Sciences at ANU, with a mission to collect and preserve Australian social science data on behalf of the social science research community
- ADA holds over 5000 datasets from around 1500 studies, including national election studies; public opinion polls; social attitudes surveys, censuses, aggregate statistics, administrative data & other sources.
- Data holdings are sourced from academic, government and private sectors.

ADA Subarchives

- 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
- Qualitative – 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



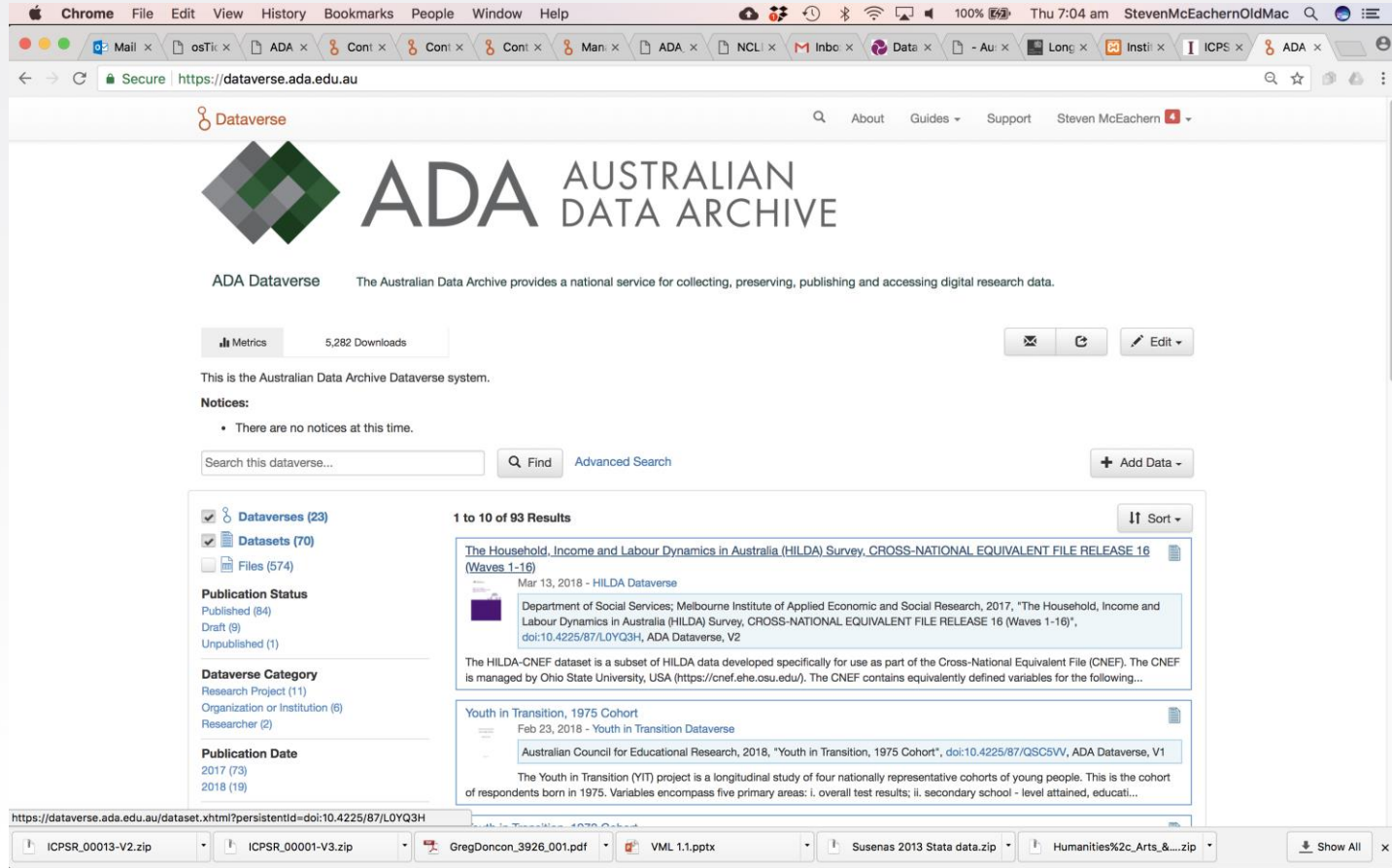
ADA Data Holdings

ADA data holdings cover a wide variety of subject areas:

- Ageing
- Business and management
- Census data
- Culture
- Demography
- Drugs, alcohol and tobacco
- Economics
- Education, employment and work
- Environment, Conservation, Land use
- Family studies
- Foreign affairs
- Gambling
- Health
- Housing
- Law, Crime, Courts
- Mass media, communication and language
- Migration, immigration and multiculturalism
- Politics and elections
- Public opinion and social attitudes
- Psychology
- Quality of life
- Science, Technology
- Social welfare
- Sociology
- Tourism, recreation and leisure
- Travel and transport

<http://dataverse.ada.edu.au>

- Redevelopment of our database and website infrastructure
 - New website
 - New data catalogue
- New functionality:
 - Self-deposit of data
 - Open data access
 - API access (both for deposit and access, e.g. through R)
 - Shibboleth authentication
 - ORCID integration
- Currently completing data migration
 - Major studies already migrated



The screenshot shows a Chrome browser window displaying the ADA Dataverse website. The browser's address bar shows the URL <https://dataverse.ada.edu.au>. The website header includes the ADA logo and the text "AUSTRALIAN DATA ARCHIVE". Below the header, the site's purpose is described: "The Australian Data Archive provides a national service for collecting, preserving, publishing and accessing digital research data." A metrics bar indicates "5,282 Downloads". A search bar is present with a "Find" button and an "Advanced Search" link. On the left, there are filters for "Dataverses (23)", "Datasets (70)", and "Files (574)", along with "Publication Status" and "Dataverse Category" options. The main content area displays search results for "1 to 10 of 93 Results". The first result is "The Household, Income and Labour Dynamics in Australia (HILDA) Survey, CROSS-NATIONAL EQUIVALENT FILE RELEASE 16 (Waves 1-16)", dated Mar 13, 2018. The second result is "Youth in Transition, 1975 Cohort", dated Feb 23, 2018. The browser's taskbar at the bottom shows several open files, including "ICPSR_00013-V2.zip", "ICPSR_00001-V3.zip", "GregDoncon_3926_001.pdf", "VML 1.1.pptx", "Susenas 2013 Stata data.zip", and "Humanities%2c_Arts_&...zip".



Platforms for Humanities, Arts AND Social Sciences?

Significant amount of crossover. For example:

- Census data
- Qualitative data
- Historical collections
- Indigenous data

Needs

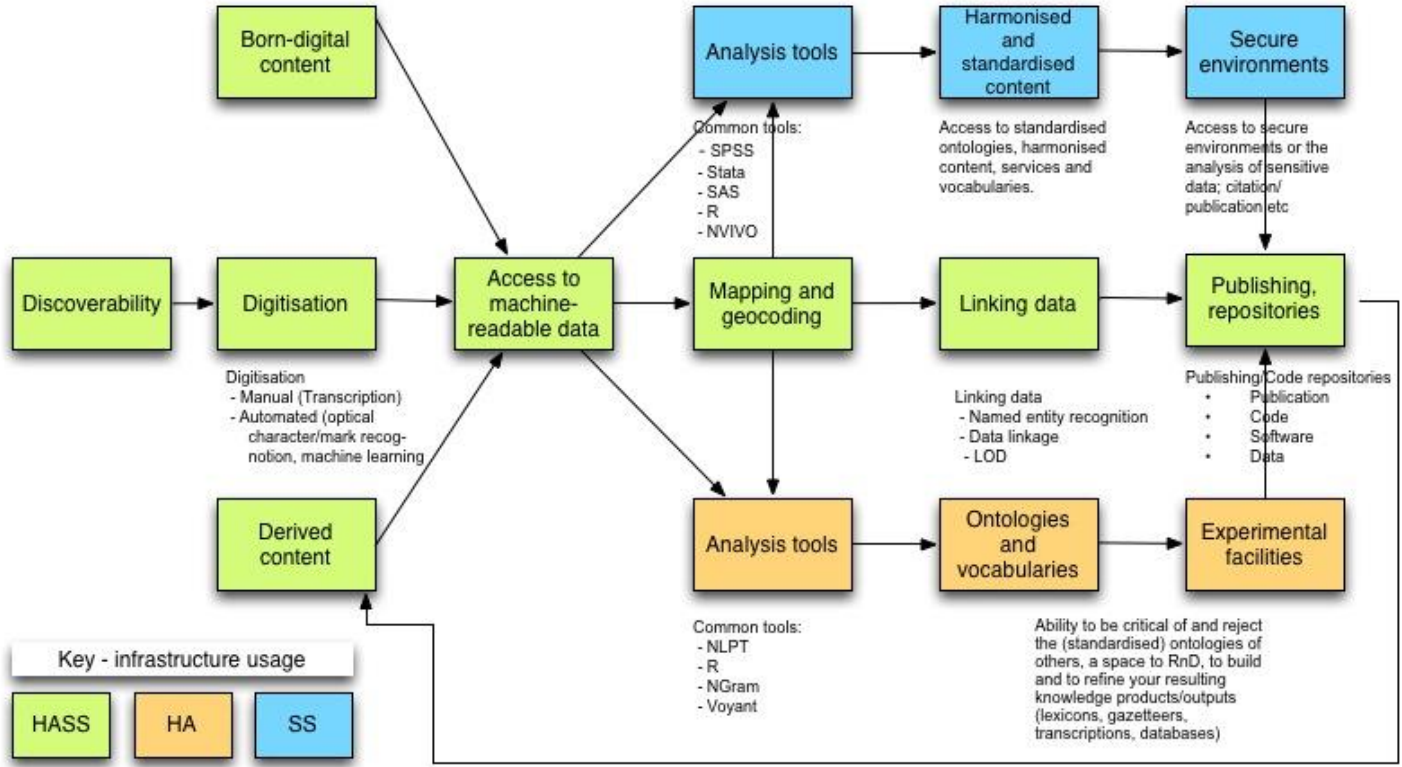
- Shared tools
- Shared data
- Shared environment

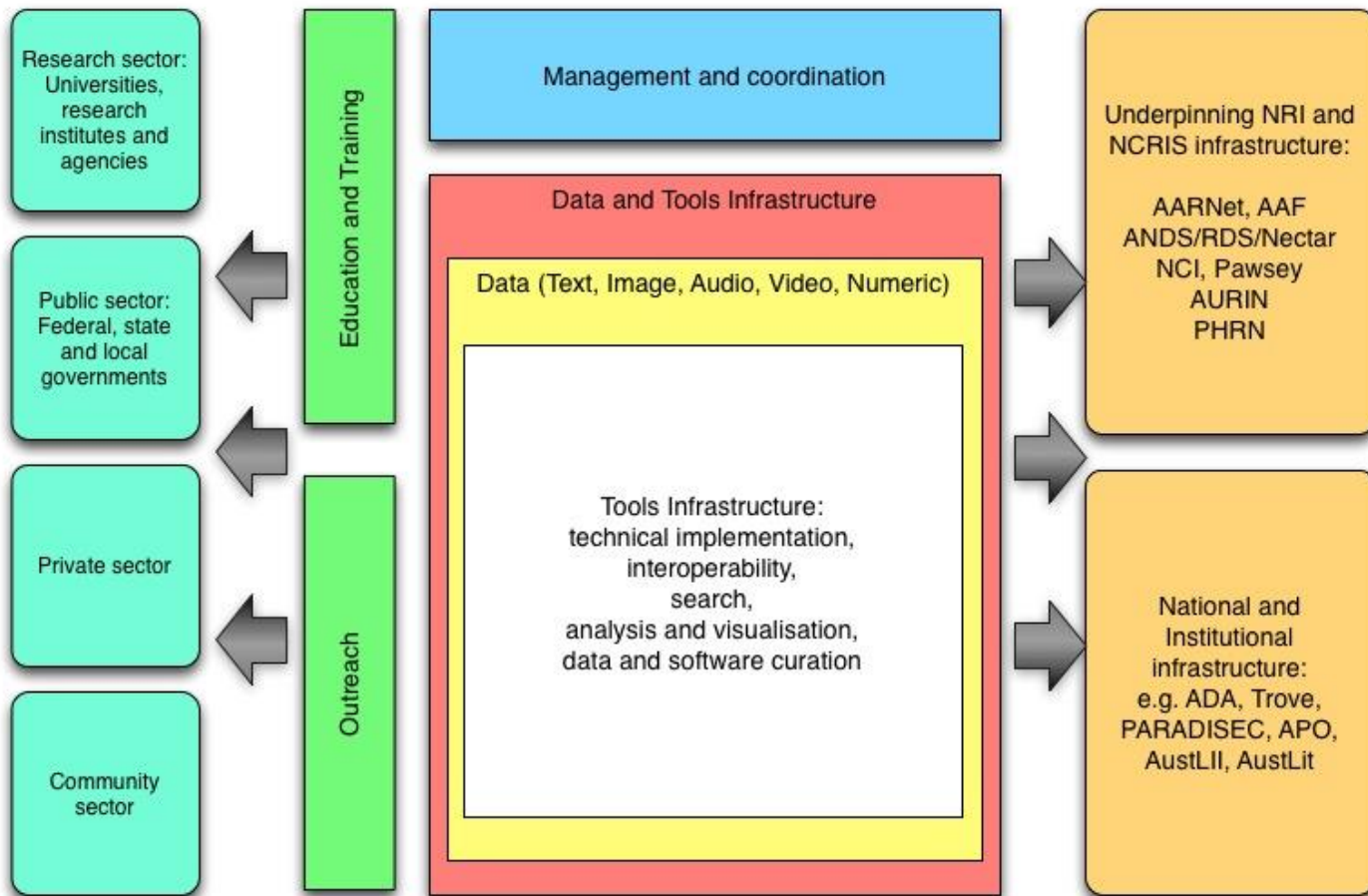
BUT ALSO

- Discipline-specific tools
- Discipline-specific methods
- Discipline-specific methods

A possible model

Social Science and Humanities data lifecycle/infrastructure requirements overlap







Questions?

Steven McEachern

steven.mceachern@anu.edu.au

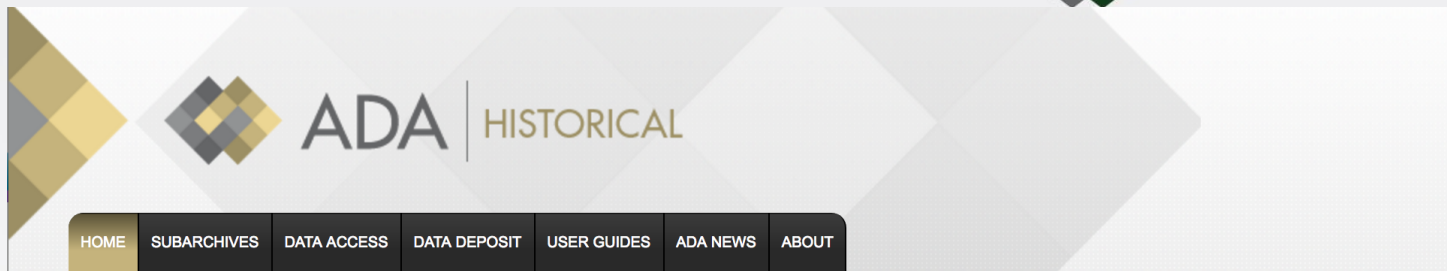
ada@anu.edu.au

<http://ada.edu.au>



HCCDA

- The Historical Census and Colonial Data Archive (HCCDA) is an archive of Australian Colonial census publications and reports. It's a sub-archive of the Australian Data Archive (ADA)
 - Collaboration between ABS archive and ADA from 2007 - 2011
- Note that the archive contains colonial census reports and not the individual census returns



HCCDA HOME DOCUMENTS TABLES ABOUT HELP CONTACTS TERMS

YOU ARE HERE: /HCCDA/PAGES/NSW-1891-CENSUS-02_3

HCCDA Document 'NSW-1891-census' page 3:

« first < previous 01_7 02_1 02_2 02_3 02_4 02_5 02_6 next > last »

CENSUS OF NEW SOUTH WALES, 1891.—AGES., TABLE II.—Showing the Number of Persons, Males and Females (exclusive of Aborigines), at, above, and under each Year of Age.

Ages.	Population at each Year of Age.			Population at and above each Year of Age.			Population under each Year of Age.		
	Males.	Females.	Total.	Males.	Females.	Total.	Males.	Females.	Total.
0	17,488	17,141	34,629	608,003	515,951	1,123,954			
1	16,380	15,724	32,104	590,515	498,810	1,089,325	17,488	17,141	34,629
2	17,055	16,423	33,478	574,135	483,086	1,057,221	33,868	32,865	66,733
3	16,520	16,321	32,841	557,080	466,663	1,023,743	50,923	49,288	100,211
4	16,197	15,745	31,942	540,560	450,342	990,902	67,443	65,609	133,052
5	15,734	15,250	30,984	524,363	434,597	958,960	83,640	81,354	164,994
6	14,988	14,759	39,747	508,629	419,347	927,976	99,374	96,604	195,978
7	14,564	14,247	28,811	493,641	404,588	898,229	114,362	111,363	225,725



A grand idea: A 250 year census time series of Australia

- The year 2020 is the 250th anniversary of Cook's landing at Botany Bay
- What if we could create a 250 year census time series of the historical population of Australia?
- We know what we have
- We know what else is needed

- What resources would be needed?
- How long would it take?

Extensions

- Vital statistics
 - Births, deaths and marriages, e.g. <https://www.bdm.vic.gov.au/research-and-family-history/apply-for-registry-data>
 - Australian Demographic Databank
- Aboriginal population
 - e.g. The Annual Aboriginal Census 1921-1944 (Smith and Briscoe)
 - Smith, L.R. (1980) The Aboriginal Population of Australia. Aborigines in Society; no. 14. Canberra: ANU Press
 - Briscoe, G. and Smith, L.R. (2001) The Aboriginal Population Revisited: 70,000 years to the present. Canberra: Aboriginal History Incorporated.
- Election results
 - <http://psephos.adam-carr.net/>
 - <http://elections.uwa.edu.au/>
- Crime
 - Australian Crime and Economic Statistics, 1900 – 1980
 - Australian Institute of Criminology - <http://aic.gov.au/statistics/homicide.html>
- Weather and climate
 - Bureau of Meteorology

Why does this matter?

- Studying long-term trends in the Australian population and related measures helps us to better understand the relationships between social, economic and environmental change
- Studying such changes requires long-term time series from multiple data sources – **including population estimates**
- e.g. relationship between drought and suicide in Australia
 - Page, A., Morrell, S. & Taylor, R. (2002). Suicide and political regime in New South Wales and Australia during the 20th century. *Journal of Epidemiology and Community Health*, 56, 766.
 - Hanigan, I.C., Butler, C.D., Kokic, P.N. & Hutchinson, M.F. (2012). Suicide and drought in New South Wales, Australia, 1970-2007. *Proceedings of the National Academy of Sciences of the United States of America*, 109, 13950-13955.