New Zealand Social Science Data Service
(www.nzsssds.org.nz)

Seminar to update on operations and future plans
Friday, 3rd September 2010
Outline

• Background
• Vision
• Origins
• Date service functions
• Live demonstration
• Issues
• Where to from here?
COMPASS structure

Advanced Network & GRID

COMPASS Research Centre
Data management & Statistical services

User Community: Researchers, Students, Government

Data Services, catalogues

NZSSDS
Data preservation, access, sharing & analysis

NZSSSN
Knowledge transfer & training
Why preserve or share data?

- Analyse data in ways not originally envisioned
- Cheaper to reuse existing data than collect new
- Provides resources for education and training
- Allows testing of new or alternative methods
- Promotes the research that created the data
- Enables scrutiny of research outcomes
- Preserves historically important data
Origins of NZSSDS

- An ambitiously titled project

  Developing a hub for e-Social Science in New Zealand: Linking the grid with emergent social data technologies and networks

- Start date 1st July 2007; project term 1 year
- Funded by the Tertiary Education Commission
Who was involved

- Set up a broadly based stakeholder group including
  - Statistics New Zealand (Stats NZ)
  - Broadband-enabled Science and Technology Grid (BeSTGRID)
  - Building Research Capability in the Social Sciences (BRCSS)
  - Data Saving and Sharing Committee, SPEaR (Ministry of Social Development)
  - Massey University
  - University of Canterbury
• Based on the Australian Social Science Data Archive (ASSDA, assda.anu.edu.au)
• ASSDA staff visited UoA to assist in initial setup and dissemination
• ASSDA provided initial data sets – they had been holding NZ data sets in the meantime
• Reciprocal visit to ASSDA covered more of the technical details
Data service functions

- Preserving and making available survey data sets & metadata
- Making available ‘Enhanced Publications’
- Making available teaching data subsets and associated workbooks
Preserving and making available survey data sets & metadata
Preserving and making available data and metadata

- Almost 50 data sets archived including
  - New Zealand pre-election surveys & other regional voting surveys (1972–1981)
  - International Social Survey Programme data, New Zealand (1991–2007) – Professor Philip Gendall
  - World Internet Project for New Zealand (2007) – Professor Charles Crothers
  - Health data sets (adverse events, oral health care, primary care, sexual health) – Professor Peter Davis
Usage history

• 50 registered users; 12 specific data requests
  – Extensive use of election study data within The University of Auckland as well as by several university students internationally
  – Use of ISSP, oral health survey and primary care survey data by researchers in Holland
• 2 postgraduate courses (in sociology & stats) involving metadata and data sets in teaching
• Further studies to follow up for archiving
Future data set developments

• Increase the number and range of data sets held on the archive
  – Initiatives in progress
    > Professor John Hattie, Faculty of Education – student assessment data
    > Commonwealth Foundation – international health survey data
    > World values study (Paul Perry – Massey University)

• Obtain high quality high value qualitative data set(s)
  > Both for preservation and for teaching purposes

• COMPASS hosting visiting speaker Louise Corti of UK data archive
Data set enhancement work plan Sep 2010 – Aug 2011

- New data sets: NZES 2008, ISSP 2008 and 2009, CoMedCa, further teaching data sets including the MoJ crime survey
- Further development of website
  - Links to Stats NZ around data access agreements with universities
  - Links to and local storage of COMPASS reports on Stats NZ website
  - Complete in-line html presentation of Census Data Guide
  - Add value to data presented on Nesstar
Enhanced publications
Enhanced publications

• What are they
  – An enhanced publication is a publication that is enriched with three categories of information:
    > research data (evidence of the research)
    > extra materials (to illustrate or clarify)
    > post-publication data (commentaries, ranking)

• Why are they useful
  – promote the availability of reusable scientific data
  – allow verification of the outcomes of the research.
  – they are the outcomes of public good research and should therefore, where appropriate, be available to the public
  – reduce the need to ‘re-invent the wheel’ and thus make better use of scarce resources
 Enhanced publication

- Derived from the work of summer students Dec 2009 Feb 2010 – project titled ‘Adding Value to the NZSSDS’
- Produced a user guide for establishing a template for a research repository with real-world examples
Divided into five sections

• Article Details
  > Background, Objective, Methods, Conclusion

• Purpose of Metadata

• Using the Data
  > Variables on the files

• Analyses

• Data Dictionary
  > Original Variables
  > Derived Variables
Teaching Resources
Teaching resources

• Why are these needed?
  – Shortage of quantitative skills among social scientists
  – Resources are expensive / time consuming to produce
  – Part of COMPASS contribution to social science community

• What are they?
  – Teaching workbooks
  – Teaching datasets
Teaching resources

• Teaching workbooks
  – Two produce so far
    > ISSP – gender roles module
    > NZES - analysis of election data
  – In the pipeline
    > MOJ Crime and Victimisation Survey

• Teaching datasets
  – Two being produced
    > ISSP
    > NZES
Demonstration

- Overview of the NESSTAR software
- Demonstration of the three functions of the data service
Nesstar software (1)

- Networked Social Science Tools and Resources
- Brings the advantages of the web to data storage
- Data archives in UK and Norway run the company
- Nesstar used by 100+ institutions, all over the world, including NZSSDS, Australia’s ASSDA
- CESSDA – 20+ member archives running Nesstar
  - Global keyword search online at cessda.org across all of these
Nesstar software (2)

Administrators
(Archivist/Data Librarian)

User community
(Researchers, teachers, students)

IT administrators and infrastructure
(technicians and architects)
Live demonstration of the three data service functions
What issues do we face?

• Financial sustainability

• Uptake of the service
  – Depositing of data
  – Use of the data archive
Where to from here?

- Continue the development of the three functions of the data service
  - Increase the number of datasets
  - Add additional enhanced publications
    > Using summer students
  - Develop additional teaching resources

- Work to advance the data preserving and sharing agenda in New Zealand
  - Data archiving on the policy agenda – see MoRST lead initiative
    ‘Data matters’ as part of the eResearch agenda
Contact details

• For further information visit
  – www.nzssds.org.nz

• Or contact
  – admin@nzssds.org.nz