Challenges and threats to data collection and usage

Presentation at Welfare Reform: Meeting the Policy Challenges of Change, 17-18 September 2015, Australian National University, Canberra

"....every major scientific revolution has been driven by one thing, and that is data.” Shaw, 2014, Harvard Magazine
Context -- contradictory imperatives

Policy
- More complex, diverse society
- Easy and rapid communication of 'information'
- 24 hour media cycle
- Multiple sources of data and interpretation
- Contestability of policy advice
- Demands from ministers
- Moral panics
- Rights based and increasingly litigious community
- Increased expectation of transparency in decision making

Academic
- ARC wont fund regular data collection, $$ per grant are low, does not cover academic salaries
- Success rates for social science competitive grants low
- No incentive for curation and depositing of data
- Rewards are weighted to 'solo' research activity
- Promotions system
- Citations - rankings
- Multiple jobs - teach, grant raising, research/publish, administration, service, community outreach
Big data

Challenges

- Identifying what matters
- Sophisticated analysis - new software tools
- Linkage of large administrative datasets
- Managing teams
- Communicating the value of data
- New frontiers - textual and visual data

Threats

- Privacy
- Access
- Failure to invest in research
- Risk over quantification
- Lack of rewards to make data 'available'
Ethics committees

Challenges

- Committee members
- Process - red tape
- Timeliness
- Managing 'consent'
- Confidentiality and linkage
- Committees being used as proxies for assessing legal or reputational risk to governing institution

Threats

- Quality of the research
- Risk averse
- 'Conservative' research
- Controls over depositing, access and preservation
Right to privacy

Challenges

- Public confidence in 'science'
- Understanding confidentiality
- Too many surveys
- Collection of repititous data
- Moving to a new discussion about the need for access to key data to inform policy making

Threats

- Public trust/legitimacy
- Stalking horse
- Lawyers
- Media
Risk averse public service

Challenges
- Dealing with the findings
- Why no effects matter
- $$$ -- good research costs
- Skills to understand methodology - poor methods training of undergraduate and PhD students
- Collaboration -- has more benefits than negatives
- Producer not just consumer of data

Threats
- Failure to collect high quality data - poor policy
- Failure to evaluate - waste of money and resources
- FOI - media always looking for a headline
- Inward looking - reaffirms the status quo
- Silos - internal competitiveness crowds out linkages
- Program goals - unrealistic and not based on evidence
- Turnover in key staff
Outsourcing

Challenges
- Managing quality control of information
- Consistency across service providers
- Who owns the data and what does that mean in practice?

Threats
- Can't access key data
- No linkage across datasets
- Consistency over time when providers change
- Competitiveness between providers
Open data and archiving

Challenge
- Loss of 'control'
- Managing 'adverse' findings
- Data 'quality'
- Documentation - common standards
- Preserve the historical record
- Publish or perish
- Qualitative and visual data
- Intellectual property - short-term versus national benefit

Threat
- Gatekeepers
- Privacy concerns
- Ethics requirements
- Cost
- Volume of data
- Restrictions on access
- Funding for archives
Advantages of open data for academics from Kingsley and Brown

- More exposure for your work
- Practitioners can apply your findings
- Higher citation rates
- Taxpayers get value for money
- Compliant with grant rules
- Your research can influence policy
- The public can access your findings

Kingsley Brown
Why does this matter?

- Quality of advice/research findings
- Quality of training and reduce wasteful data collection
- Enhance quality of underlying data
- Replication
- Ability to conduct evaluations - linked datasets
- Digital record for historical purposes
- Value for money and new analysis
- Build a strong system to support domestic research and its uptake
- Innovation and global competitiveness - serendipity and business
Open inquiry is at the heart of the scientific enterprise. Publication of scientific theories - and of the experimental and observational data on which they are based - permits others to identify errors, to support, reject or refine theories and to reuse data for further understanding and knowledge. Science’s powerful capacity for self-correction comes from this openness to scrutiny and challenge.