Data for Science: where to from here?

Warwick Cathro
Assistant Director-General, Innovation,
National Library of Australia

NATIONAL
LIBRARY OF AUSTRALIA

Background on PMSEIC

- Prime Minister's Science Engineering and Innovation Council
- Direct source of advice to the Government
- 10 senior Ministers, 20 others
- Working groups

NATIONAL LIBRARY OF AUSTRALIA

Working Group on Data for Science

- Formed in April 2006
- Terms of Reference (summary):
  - Outline issues concerning management of scientific and research data
  - Identify data management and data access strategies
- Co-chaired by Robin Batterham and Fiona Stanley

NATIONAL LIBRARY OF AUSTRALIA

Working Group process

- Met monthly May-October 2006
- Secretariat support from Office of Chief Scientist
- Presentation to PMSEIC, 8 December

NATIONAL LIBRARY OF AUSTRALIA

Information gathering

- Submissions (46)
- Context:
  - In Australia: SII, NCRIS, e-RCC
  - Buchhorn/McNamara report
  - Overseas, ICSU, UK e-Science Initiative, US National Science Board, etc.
- Case studies:
  - Global Initiative on Sharing Avian Influenza Data
  - Australian Ocean Data Centre Joint Facility
  - PARADISEC as a vulnerable dataset

NATIONAL LIBRARY OF AUSTRALIA

Issues considered

- Volume of data
- Value of data for national problem solving
- Changing research methods
- Stewardship issues
- Privacy issues

NATIONAL
LIBRARY OF AUSTRALIA

Summary of findings

- A national framework
- A national network of digital repositories
- Data management plans
- Measures to make data more accessible
- Skills development
- Debate on the “national centre”

NATIONAL LIBRARY OF AUSTRALIA

A national framework

- There needs to be a national approach for managing Australia's research data
  - "a nationally supported long term strategic framework … including guiding principles, policies, best practices and infrastructure" (Rec.1)
  - "high level expert committee" (Rec.2)

NATIONAL LIBRARY OF AUSTRALIA

A national network of repositories

- Put in place the policy and programmes needed to establish a “sustainable publicly funded national network of federated digital repositories” (Rec.3,4)
- Intended to cover both institutional and collaborative data centres

NATIONAL LIBRARY OF AUSTRALIA

Data management plans

- Sustainability needs to start with the researcher who generates the data
- Noted the recommendations of the US National Science Board
- Funding agencies should “offer incentives to encourage researchers and institutions to develop data management plans” (Rec.7)

NATIONAL LIBRARY OF AUSTRALIA

Making data more accessible

- “Open equitable access” as a general principle (Rec.6)
- Ensure we implement standards to support inter-operability (Rec.5)
- Remove “regulatory impediments” to data sharing
  - unnecessary IP barriers to data sharing (Rec.9)
  - balance on privacy issues (Rec.8)

NATIONAL LIBRARY OF AUSTRALIA

Skills development

- Also highlighted by other reports:
  - the US National Science Board
  - the e-Research Coordinating Committee
- Measures to ensure that “data management expertise becomes a core skill for researchers” (Rec.10)
- Supported findings of e-RCC (Rec.11)

NATIONAL LIBRARY OF AUSTRALIA

A National Centre?

- Debated but not agreed
- Roles for a Centre may have included:
  - Designation of data archives
  - Recommendations on funding of data archives
  - Support for researchers (eg with data management plans)
  - Support for standards development
  - Policy coordination (eg access protocols)
  - Support for data archives (eg advice on sustainability)

NATIONAL LIBRARY OF AUSTRALIA

Reflections on the process

- Meetings
- Office of Chief Scientist
- Cabinet Room, 8 December 2006

NATIONAL
LIBRARY OF AUSTRALIA