



# Morning Tea Break

Please be back by 11:15am



## More about the *Code*

# What does the *Code* say about data? (1)

- ‘The responsible conduct of research includes the **proper management and retention** of the research data.’

## What does the *Code* say about data? (2)

- ‘Each institution must have a policy on the retention of materials and research data.’
- ‘Institutions must provide facilities for the safe and secure storage of research data and for maintaining records of where research data are stored.’

# Responsibilities of institutions

- Retain research data and primary materials
- Provide secure research data storage and record-keeping facilities
- Identify ownership of research data and primary materials
- Ensure security and confidentiality of research data and primary materials

## What does the *Code* say about data? (3)

- ‘The researcher must decide **which data and materials should be retained**, although in some cases this is determined by law, funding agency, publisher or by convention in the discipline. The central aim is that sufficient materials and data are retained to **justify the outcomes** of the research and to **defend them if they are challenged**. The potential **value of the material for further research** should also be considered, particularly where the research would be difficult or impossible to repeat.’

# Responsibilities of researchers (1)

- Retain research data and primary materials
  - As long as interest and discussion persist
  - Make it available to others
  - At least as long as the institution requires
  - If research is challenged, everything must be kept until the challenge is resolved
  - If there is an allegation of misconduct, everything must be kept
  - Secure and safe disposal

## Responsibilities of researchers (2)

- Manage storage of research data and primary materials
  - Keep clear and accurate records
  - Safe and secure storage
  - Keep primary records too (e.g. lab notebooks)
  - Keep data in durable, indexed and retrievable form
  - Keep a catalogue
  - Follow ethical protocols and legislation

## Responsibilities of researchers (3)

- Maintain confidentiality of research data and primary materials
  - Only use confidential data in ways agreed with providers
  - Take particular care when making confidential data available for discussion

# Tension?

- Possible tension between the institution and the researcher
- Differences in interpretation of division of responsibilities
- Implications of centralised vs. devolved responsibility/power etc?

Any questions?



# Institutional responses to the *Code* and strategies for implementation

# Implications for institutions

- Data must be:
  - 'Owned'
  - Retained
  - Stored
  - Retrievable
  - Accessible
- Institutions need to provide the policies, infrastructure, support services and facilities to do this
- Researchers cannot meet their obligations without institutional support

# Key considerations

- Recognising that the *Code* is about cultural change as well as research practice
- Ensuring an appropriate governance framework with representation from all interested parts of the institution — research office, library, IT, archives & researchers
- Ensuring buy-in from senior administrators
- Providing both data management infrastructure and services to support compliance

## The Challenge

- Effective collaboration between these services and areas
- Getting the institution to accept the need for it
- Using or adapting existing policy and services to enable compliance

## Getting started

- High level support for the work – who is the ‘Champion’?
- Steering process to coordinate work
- Identification and review of what policy and infrastructure is already in place
- Identification of resources required to provide support – this will vary between institutions

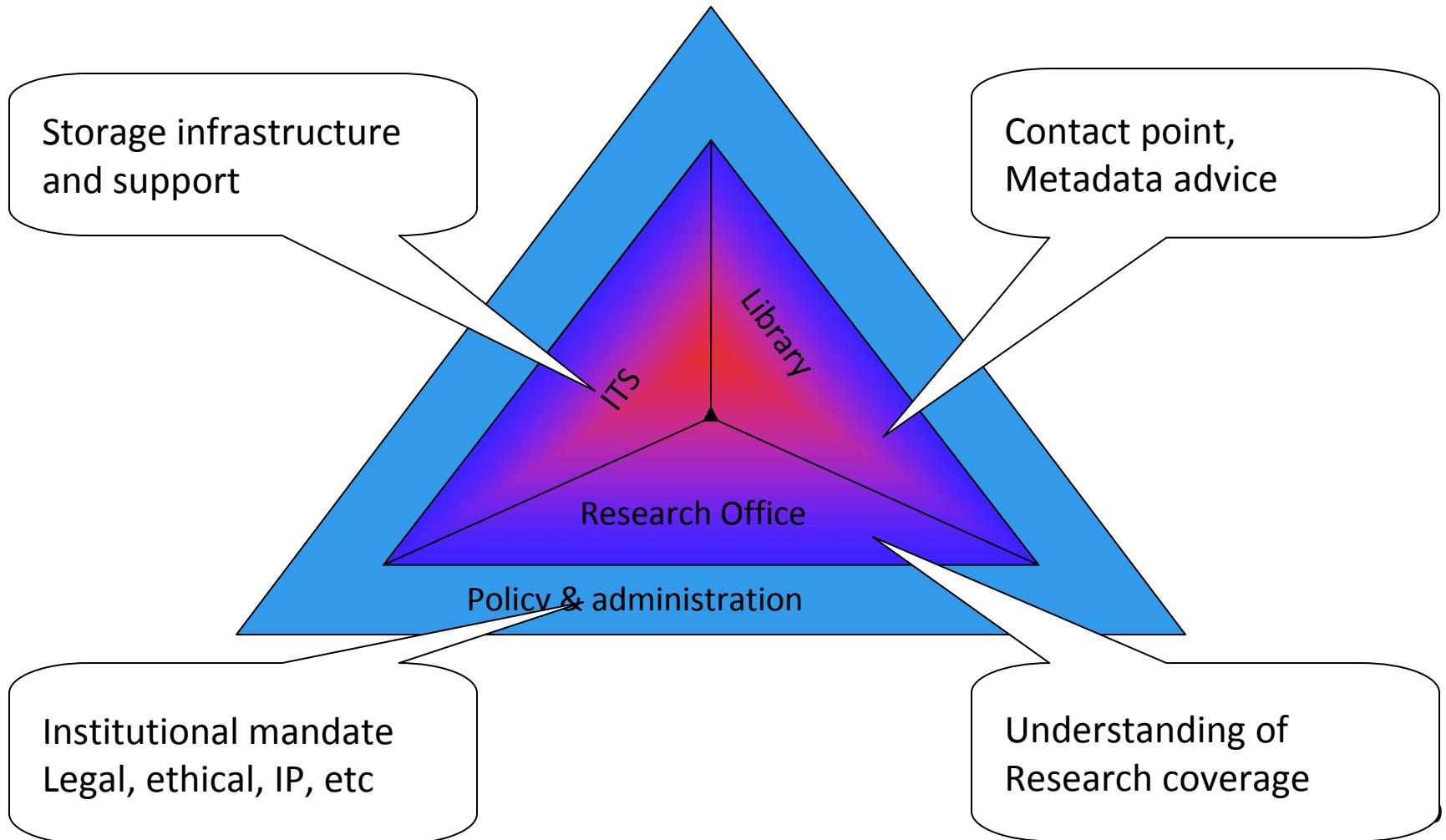
## First steps in responding to the *Code* (1)

- Activities for assessing the readiness and capability of the institution (as opposed to the individual researcher):
  - an assessment of institutional commitment and priorities
  - a review of all existing relevant policies and responsibilities, especially information management policies, practices, planning and responsibilities

## First steps in responding to the *Code* (2)

- a review of data storage capacity, including possible use of cloud services
- an assessment of capability of individual researchers to respond to the *Code*, through, for example, a survey of data management practices
- an audit of data currently held within the institution
- consideration of the financial implications

# Institutional support areas



## Potential role of Research Office

- Policy creation & oversight
- Identifying significant researchers and data producers
- Connecting data with publications
- Leadership and advocacy role
- Research practice training and oversight

## Potential roles of library/archives/records

- Contact and outreach to researchers
- Information gathering about available data
- Advice on metadata, descriptions, disposal policy and sustainability
- Possible use of institutional repository for holding descriptions or connecting publications to data
- Training and support

## Potential role of ITS/ICT

- Provider and supporter of storage hardware, back-ups
- Provider and supporter of tools for data capture, metadata capture and discovery
- Access control
- Connections to the ARCS Data Fabric

## Policy and administration

- Institutions already have policies and procedures covering research, records management, intellectual property and other research-related activities.
- These can be adapted for *Code* compliance purposes.

## Related policies (1)

1. Intellectual property — covering copyright, moral rights, patents
2. Collaboration and contractual agreements — do these include provision for data sharing after the research is complete and agreement on who will host and store the data
3. Conflict of interest — does this impact on ability to make data available

## Related policies (2)

### 4. Data management, including:

- Storage — how to provide appropriate storage and under what conditions
- Retention — how long data is to be stored
- Disposal — how this is to be recorded and managed
- Access — how to make data available to those engaged in the research project; how data is to be “published” and made available more widely; how the institution is to keep a record of its data assets; the institutional position on open access of both data and publications

## Related policies (3)

5. Ethics and privacy — how these affect the length of data storage and the ability to share
6. Compliance — what measures are in place to track this.

Each of these needs to be considered in terms of both institutional and researcher responsibilities.

Not all are specific to data management, but each should have a data management component and requires a procedure translating policy into practice.

## The Strategic Response to the *Code*

- Once an appropriate policy framework is in place, there will be a need to develop strategies to cover:
  - publicising the policies within the community
  - providing appropriate infrastructure to support researcher requirements and training in data management and use of infrastructure
  - establishing support services related to the infrastructure
  - establishing record-keeping procedures
  - demonstrating compliance through review or audit frameworks.

## Costs of *Code* compliance

- Policy review and amendment
- Legal review and changes
- Potential view as a burden
- More work for existing services
- Creation of compliance monitoring
- IT infrastructure

## Benefits of *Code* compliance

- Security of storage
- Enhanced ability to share (including “sharing with yourself”)
- Better understanding of what your researchers are doing
- Establishing a corpus of past research to build upon
- Readiness if the *Code* does become mandatory
- Readiness as funders start to require it

## Key messages

- If it is not easy and clear, researchers won't/can't do it
- There are benefits for the organisation in doing this
- This process is about change and it will be gradual
- Effective data management comes from collaboration across the organisation
- ANDS can help

Any questions?



# Data sharing and re-use

# Data sharing and re-use

## What the *Code* has to say

- ‘A strong research culture will demonstrate ... good stewardship of public resources used to conduct research.’
- ‘The potential value of the material for further research should [...] be considered, particularly where the research would be difficult or impossible to repeat.’
- ‘Research data should be made available for use by other researchers unless this is prevented by ethical, privacy or confidentiality matters.’

The ANDS position: data should be made available unless there is a good reason not to.

## *Report of the Government 2.0 Taskforce*

- public sector information is a national resource and that releasing as much of it on as permissive terms as possible will maximise its economic and social value to Australians and reinforce its contribution to a healthy democracy

## *Report of the Government 2.0 Taskforce (2)*

Recommendation 6: Make public sector information open, accessible and Reusable

- 6.1 By default Public Sector Information<sup>19</sup> (PSI) should be:
  - free
  - based on open standards
  - easily discoverable
  - understandable
  - machine-readable
  - freely reusable and transformable.
- 6.3 Consistent with the need for free and open re-use and adaptation, PSI released should be licensed under the Creative Commons BY standard<sup>24</sup> as the default.
- And more ...

# Attitudes to data sharing

- 8.6% share openly
- 44.0% share via negotiated access
- 6.4% share after the formal end of a project
- 19.0% don't share at all
- 17.6% don't share because of legal and confidentiality issues
- 13.13% don't share but would be willing to if there was an easy means of doing so

*Investigating Data Management in Australian Universities*, by Margaret Henty, Belinda Weaver, Stephanie Bradbury and Simon Porter, APSR, 2008

[www.apsr.edu.au](http://www.apsr.edu.au)

## Data sharing and re-use – the practicalities

- How will the data be made available?
- Where will the data be stored?
- How will the dataset and its accompanying documentation be described?
- Who will provide that description?
- Will the data be openly available or available only on request?
- Are there embargo periods?

# Benefits of data sharing for the researcher

- Greater recognition of their work through increased citations and follow-on research
- Multiple copies protect the data from accidental erasure
- May lead to new collaborations with data users
- Demonstrates the value of your work by allowing continued re-use of the data, which in turn may lead to further funding
- Meets the contractual obligations of research funders who have data sharing policies
- Leads to satisfaction as a result of benefiting the broader community

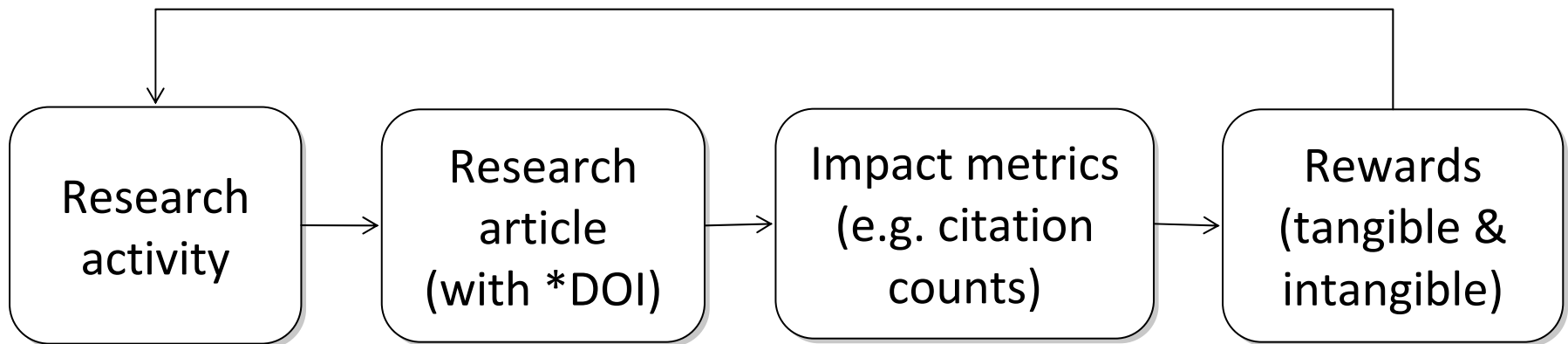
## Broader benefits of data sharing

- Productivity increases through reduction of costs.
- A community of researchers can develop around a common resource.
- Independent lines of research based on the same data might result in a better research outcome.
- The risk of duplicating data collection efforts is reduced.

# Obstacles to data sharing

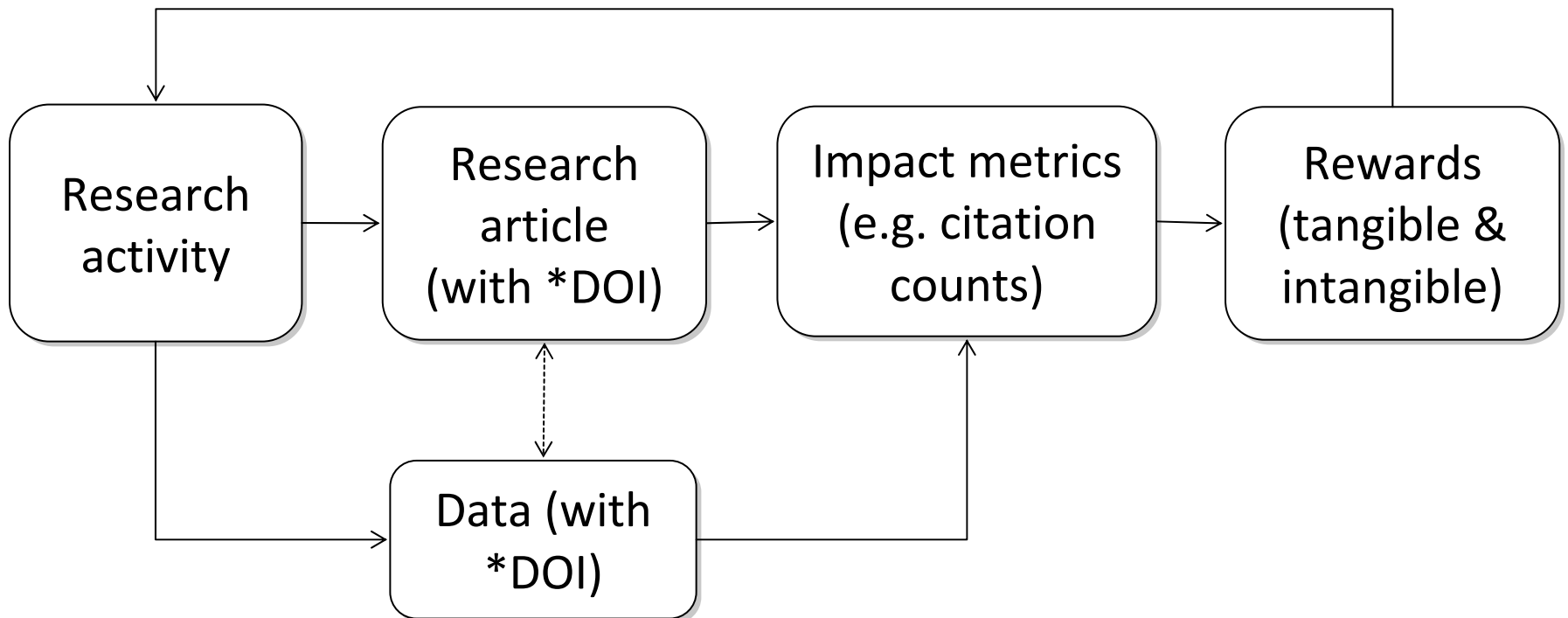
- Lack of institutional policies and frameworks to support data sharing
- Not yet finished publishing
- Lack of time
- Lack of resources
- Confidentiality
- Lack of reward

## Rewards: what we have...



\*DOI: Digital Object Identifier

# Rewards: what we need...



\*DOI: Digital Object Identifier

Any questions?



# Lunch Break

Please be back by 1:30pm