

Legal Perspectives on the Ownership and Policy Framework for Research Data

Dr Anne Fitzgerald
QUT

APSR: Long-lived Collections – the future of
Australian research data

Canberra, 11 September 2007

Legal Perspectives on the Ownership and Policy Framework for Research Data

We are going to be deluged with data in almost every field.

Tony Hey, 2006



**BUILDING THE
INFRASTRUCTURE
FOR DATA ACCESS
AND REUSE IN
COLLABORATIVE
RESEARCH:**

**AN ANALYSIS OF
THE LEGAL CONTEXT.**

Dr Anne Fitzgerald and Kylie Pappalardo.
Assisted by Professor Brian Fitzgerald,
Anthony Austin, Dr John Abbott,
Brendan Cosman, Damien O'Brien and
Bill Singleton (Allens Arthur Robinson).

Legal Perspectives on the Ownership and Policy Framework for Research Data

- ***BUILDING THE INFRASTRUCTURE FOR DATA ACCESS AND REUSE IN COLLABORATIVE RESEARCH: An Analysis of the Legal Context***

Dr Anne Fitzgerald, Kylie Pappalardo and others (June 2007)

<http://www.oaklaw.qut.edu.au/node/33>

- A joint report by QUT's OAK Law Project (Open Access to Knowledge) and the Legal Framework for e-Research Project,
- Both projects funded by the Australian Government Department of Education Science and Training (DEST)
- Systemic Infrastructure Initiative (SII) funded projects and part of the Commonwealth Government's *Backing Australia's Ability – An Innovation Action Plan for the Future*

Building the Infrastructure for Data Access and Reuse in Collaborative Research – the legal context

- Examines the legal framework in which research data is generated, managed, disseminated, used (chapter 1)
- Overviews laws relevant to research data:
 - copyright, contract, confidentiality, privacy, public records, specific laws applying to certain data collections, FOI, etc (chapters 2, 3, 5, 6, 7)
- Explains current data sharing practices and attitudes (chapter 4)
- Considers open access policies developed at national and international levels (chapter 8)
- Provides practical guidance and recommendations on the development of appropriate legal frameworks for data management (chapters 9 and 10)

<http://www.oaklaw.qut.edu.au/node/33>

<http://www.e-research.law.qut.edu.au/node/14>

Chapter 1 – The Data Landscape

“Today’s research community must assume responsibility for building a robust data and information infrastructure for the future.” [ICSU, *Scientific Data and Information: A Report of the CSPR Assessment Panel*, December 2004]

- Australian Research Council (ARC) and National Health and Medical Research Council (NHMRC) announced in December 2006 that open access guidelines for data from funded research projects would take effect from 2008
 - Approach was supported by the Productivity Commission in *Public Funding for Science and Innovation* (2007)
- Prime Minister’s Science, Engineering and Innovation Council (PMSEIC), Working Group on Data for Science, *From Data to Wisdom: Pathways for Successful Data Management for Australian Science* (December 2006)
- Australian Government National Collaborative Research Infrastructure Strategy (NCRIS), *Strategic Roadmap* (2006)
- Australian Partnership for Sustainable Repositories (APSR), *Sustainable Paths for Data-Intensive Research Communities at the University of Melbourne* (August 2006)
- NHMRC/AVCC, *Joint Statement and Guidelines on Research Practice – Australian Code for Responsible Conduct of Research* (2006)

<http://www.oaklaw.qut.edu.au/node/33>

<http://www.e-research.law.qut.edu.au/node/14>

Chapter 2 – Key Concepts

- Key terms and concepts:
 - “data”, “dataset”, “database”
 - “data” includes raw information, text, software, numbers, graphs, pictures, metadata and audio and video recordings
 - “ownership”, “control”, “access” and “use”
- Different parties may claim an interest in relation to data:
 - Researchers who generate or collect data, research funders, database managers, custodians and users
- Identifies some of the extensive range of laws that impact on issues of data ownership, control, access and use:
 - copyright, patent, contract, confidential information, privacy, statutory regulation of data collection

<http://www.oaklaw.qut.edu.au/node/33>

<http://www.e-research.law.qut.edu.au/node/14>

Chapter 3 – The Regulatory Context

“On the one hand information wants to be expensive, because it’s so valuable...On the other hand, information wants to be free, because the cost of getting it out is getting lower and lower all the time.”

[Stewart Brand, Hackers’ Conference in 1984, as quoted by John Perry Barlow in *The Economy of Ideas: A framework for patents and copyrights in the Digital Age (Everything you know about intellectual property is wrong)* (1994), Wired, Issue 2.03]

<http://www.oaklaw.qut.edu.au/node/33>

<http://www.e-research.law.qut.edu.au/node/14>

Chapter 3 – The Regulatory Context

- Examines the legal context in which data is collected, generated, managed and used
- Particular focus on legislative and administrative provisions applying to data generated or held by public sector or publicly funded entities
 - Specific legislation and regulations governing collection, handling etc of particular kinds of data
 - public records and archives legislation
 - FOI legislation
 - Electronic Transactions Acts
 - Administrative arrangements and Information Standards
- Health information and personally identifying information – legal frameworks for information privacy
 - Competing considerations: advancement of medical research (aided by data sharing and data linking practices) vs. protection of patient anonymity and privacy
 - Note Australian Law Reform Commission's inquiry into privacy – Issues Paper 31: Review of Privacy (2006) and Discussion Paper released September 2007
- Cultural protocols

<http://www.oaklaw.qut.edu.au/node/33>

<http://www.e-research.law.qut.edu.au/node/14>

Chapter 4 – Current Practices and Attitudes to Data Sharing

- Surveys and describes examples of web-accessible United States, European and Australian databases with developed legal arrangements for providing access to and sharing of research data:
 - analyses the various models of data ownership, control, access and use observed in the sample
 - looks at general frameworks as well as subject-specific databases, especially collections of medical and genetic research data

<http://www.oaklaw.qut.edu.au/node/33>

<http://www.e-research.law.qut.edu.au/node/14>

Chapter 4 – Current Practices and Attitudes to Data Sharing

Key databases examined include:

United States –

- National Institutes of Health (NIH) databases, including GenBank, Genetic Association Information Network (GAIN) and National Human Genome Research Institute (NHGRI)
- National Centre for Biotechnology Information (NCBI) databases, including Database of Genotype and Phenotype (dbGaP)
- Earth System Grid

Australia –

- Marine Themes
- Household Income and Labour Dynamics in Australia Survey (HILDA)
- Pacific and Regional Archive for Digital Sources in Endangered Cultures (PARADISEC)
- Australian Social Science Data Archive (ASSDA)
- Western Australian Genetic Epidemiology Resource (WAGER)

<http://www.oaklaw.qut.edu.au/node/33>

<http://www.e-research.law.qut.edu.au/node/14>

Chapter 5 - Copyright

- Overview of basic principles of copyright law in relation to data/datasets/databases:
 - copyright in compilations (“literary works” heading)
 - In Australia: *Desktop Marketing System Pty Ltd v Telstra Corporation Pty Ltd* (2002, Full Federal Court) – cf United States position: *Feist Publications Inc v Rural Telephone Service Co Inc* (1991)
 - European Union Database Directive
- Extension of copyright protection for digital materials available online:
 - Electronic communication right (transmit/make available)
 - Protection for Digital Rights Management (DRM) technologies and Electronic Rights Management Information (ERMI)
- Copyright licensing, including:
 - Open content licences – Creative Commons, Science Commons
 - Statutory licences – position of government as owner and user of copyright (“Crown copyright”)

<http://www.oaklaw.qut.edu.au/node/33>

<http://www.e-research.law.qut.edu.au/node/14>

Chapter 6 – Confidential Information

- Explains when data or information which is confidential or secret can be legally protected
- Doctrine of confidentiality (action for breach of confidence) can be used to control and limit access to data that has not been publicly disseminated and is not public knowledge
- Considers some special cases:
 - information acquired during the course of employment (trade secrets)
 - information held by government
- Use of contracts to maintain secrecy (“Confidentiality agreements”)

<http://www.oaklaw.qut.edu.au/node/33>

<http://www.e-research.law.qut.edu.au/node/14>

Chapter 7 - Contract

- Considers various contractual arrangements relevant to the protection and sharing of data and information
- Use of contracts:
 - to protect confidentiality or commercial interests
 - assign copyright
 - grant copyright licences
 - control access to and use of data
 - impose restrictions on re-use of data for specific purposes or projects (downstream uses)
- provides diagrammatic examples to illustrate contractual arrangements

<http://www.oaklaw.qut.edu.au/node/33>

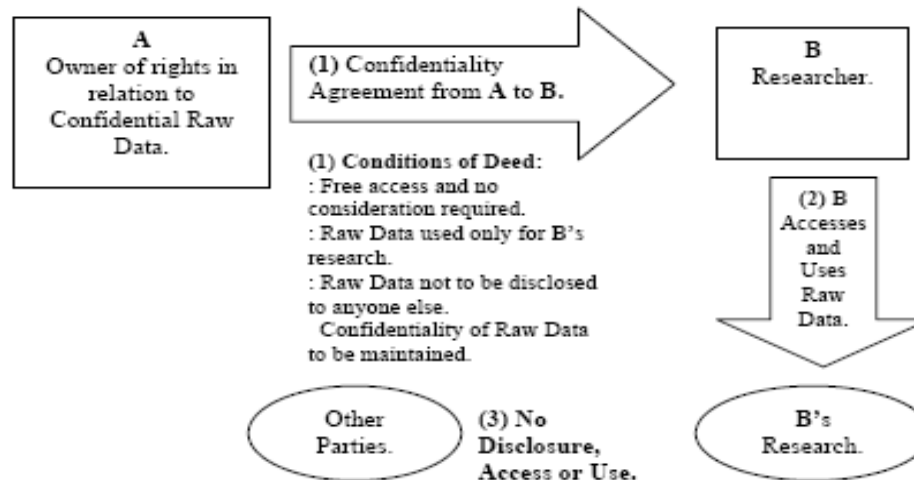
<http://www.e-research.law.qut.edu.au/node/14>

Example 1: Basic Confidentiality Agreement for Use of Research Data

Parties and Roles:

- A is the owner of rights in relation to a confidential collection of raw data. The raw data is not protected by copyright, because it does not meet the originality threshold required for copyright protection to apply.
- B is a researcher who wishes to use A's raw data in B's research.
- A and B agree to let B have access to A's data without charge, on the condition that confidentiality is maintained.

Basic Confidentiality Agreement for Use of Research Data Diagram:



Example 1 - Steps Taken:

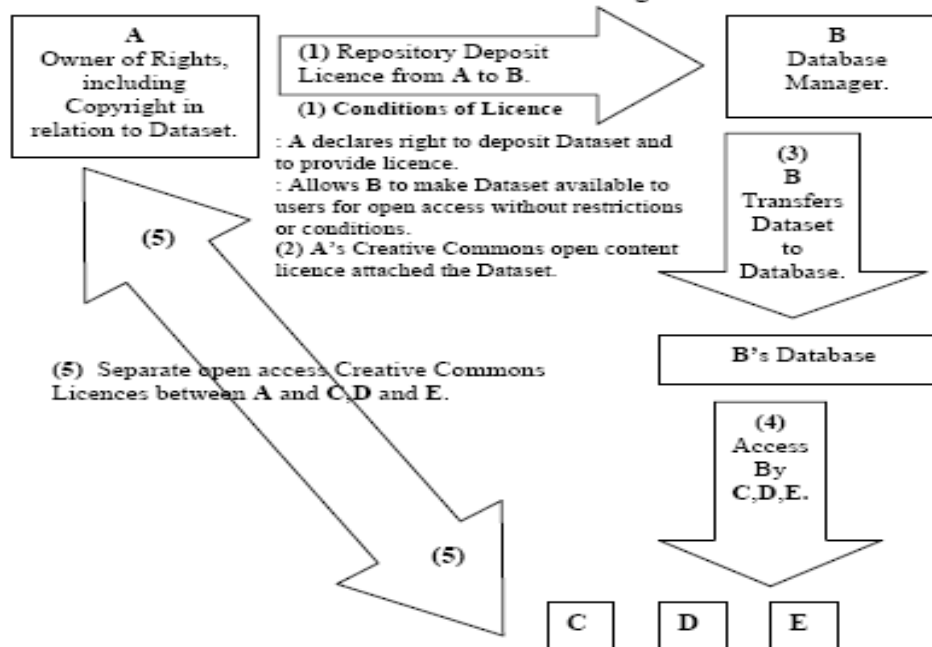
- (1) A grants B the right under a Confidentiality Agreement to use the raw data in B's research. A's access and use is free, the confidentiality agreement is made as an enforceable deed. The conditions are that B only uses the raw data for certain research purposes, that B does not disclose the raw data to anyone else and that B keeps the raw data confidential.
- (2) B accesses and uses A's raw data in B's research and ensures that there is no disclosure, access to or use of the data by other parties (3).

Example 4: Access Content Licence - Owner and User

Parties and Roles:

- A is the owner of rights, including copyright, in relation to a dataset.
- B is the managing entity of an open access database which is accessible by users on a worldwide basis without charge. B is not obliged to maintain the confidentiality of submitted data or to enter into licences with users for their access and use of the data.
- C, D and E are members of the research community who wish to access A's dataset from B's database and to use it for each of their own research projects.
- A agrees that A's dataset will be deposited into B's database and reproduced on the database for access by any users. A does not impose any restrictions or conditions on access and use of the data, other than those contained in a Creative Commons open content licence attached to the dataset by A before providing it to B.

Open Access Content Licence – Owner and User Diagram:



Chapter 8 – Data Sharing Frameworks

- Provides an overview of open access policies and principles specifically relating to research data, developed at national and international levels and by private research organisations
- International:
 - Bermuda Principles
 - Berlin Declaration on Open Access to Knowledge in Science and the Humanities
 - Organisation for Economic Co-operation and Development (OECD) Declaration on Access to Research Data from Public Funding
- National - governments and public sector research funding bodies:
 - National Institutes of Health (NIH) Data Sharing Policy; and
 - Australian Research Council (ARC) and National Health and Medical Research Council (NHMRC) funding policies.
- Private sector research organisations:
 - Wellcome Trust Position Statement (UK)

<http://www.oaklaw.qut.edu.au/node/33>

<http://www.e-research.law.qut.edu.au/node/14>

Chapter 9 – Data Sharing Infrastructure

“The scientific community can and must assert greater control over the management of its own data supplies.” [J Reichmann and P Uhler, *A Contractually Reconstructed Research Commons for Scientific Data in a Highly Protectionist Intellectual Property Environment* (2003) 66 *Law and Contemporary Problems* 315]

- Considers steps to take in developing legal frameworks for data management in the Australian research context
- In developing effective frameworks for the management of research data, the following four steps are proposed:
 - (1) formulate a Data Access Policy and Principles;
 - (2) identify the specific kinds of data to be made available for access and use;
 - (3) ascertain the conditions of access and use; and
 - (4) adopt mechanisms for practical data management, such as a Data Management Plan (DMP) and a Data Management Toolkit (DMT)

<http://www.oaklaw.qut.edu.au/node/33>

<http://www.e-research.law.qut.edu.au/node/14>

Chapter 9 – Data Sharing Infrastructure

Data Management Plan (DMP) –

- how data is managed in accordance with relevant legal controls
- how data is preserved and maintained and how this is funded
- who is responsible for the management of data and the database

Data Management Toolkit (DMT) –

- guides researchers in the allocation and handling of data in accordance with the DMP
- how to deposit data into the database
- how to manage data in compliance with legal requirements

<http://www.oaklaw.qut.edu.au/node/33>

<http://www.e-research.law.qut.edu.au/node/14>

e-Research Survey

General comment from participant:

“The problem with this issue [e-Research] is often not the desire to make data available, but the cost and difficulty in actually implementing the systems, particularly when building the systems requires people with an understanding not only of very complex technical issues for managing the data, but also the IP consequences of their actions and the possibility of legal risk (through unforeseen uses) that may accrue from release of the data....”

e-Research Survey

'How-to' guide

- Approx 90% of the sample said it would assist them to have access to a plain English 'how-to' guide explaining the legal restrictions associated with databases
- Of those who said a 'how-to' guide would NOT assist, reasons included:
 - *“...because they are already provided by the databases”*
 - *“most databases I use have no restrictions”*
 - *it “is likely to be a large document”*
 - *“I don't have time to read yet more documentation written in general terms that wouldn't tell me what I needed to know about my specific situation”*
 - *“it is the responsibility of the research office”*