

Proceedings of the
Regional Symposium on
Open Educational Resources:
An Asian Perspective on Policy and Practices

19th - 21st September 2012

This Symposium is dedicated to the International Development and Research Centre [IDRC].

 Sustainability  Quality  Policy  Technology  Practice



Main Organisers



Partners





Except where otherwise noted, the content of this proceedings is licensed under the Creative Commons Attribution 3.0 License.

Disclaimer:

The content of this proceedings is based on camera-ready papers received by 11th September 2012. The accuracy of the content and language used in the full papers published in this proceedings is the onus of the author/co-authors.

ENABLING ACCESS TO AND RE-USE OF PUBLICLY FUNDED RESEARCH DATA AS OPEN EDUCATIONAL RESOURCES: A STRATEGY FOR OVERCOMING THE LEGAL BARRIERS TO DATA ACCESS AND RE-USE

Professor Anne Fitzgerald
Law Faculty
Queensland University of Technology
Brisbane, Australia
am.fitzgerald@qut.edu.au

Haswira Nor Mohamad Hashim
Law and Justice Research Centre
Queensland University of Technology
Brisbane, Australia
haswira.mohamadhashim@student.qut.edu.au

ABSTRACT

Open Educational Resources (OER) are teaching, learning and research materials that have been released under an open licence that permits online access and re-use by others. The *2012 Paris OER Declaration* encourages the open licensing of educational materials produced with public funds. Digital data and data sets produced as a result of scientific and non-scientific research are an increasingly important category of educational materials. This paper discusses the legal challenges presented when publicly funded research data is made available as OER, arising from intellectual property rights, confidentiality and information privacy laws, and the lack of a legal duty to ensure data quality. If these legal challenges are not understood, addressed and effectively managed, they may impede and restrict access to and re-use of research data. This paper identifies some of the legal challenges that need to be addressed and describes 10 proposed best practices which are recommended for adoption to so that publicly funded research data can be made available for access and re-use as OER.

Keywords: Access, Re-Use, Research Data, Public Funds, OER, Legal Challenges

1. PUBLICLY FUNDED RESEARCH DATA AS OER

“Publicly funded research” is research conducted by government agencies or departments, as well as by other parties using public funds provided by any level of government.¹ Research data exist in both digital and non-digital formats comprising textual records, numerical scores, compilation, images, sounds and algorithms.²

There are strong justifications for enabling publicly funded research data to be made available as OER. Research activities are undertaken with the aim of creating, advancing and increasing the stock of knowledge. Dissemination and utilisation of research outputs enhances the returns on public investment in research to the society at large.³ Socio-economic benefits are gained by enabling access to and re-use of publicly funded research data as OER since economic growth in the era of the-knowledge based economy requires interactive and open dissemination of data and information.⁴ Enabling access to and re-use of publicly funded research data as

OER can stimulate innovation⁵ which draws upon existing knowledge and information.⁶

2. LEGAL CHALLENGES

There is a myriad of legal challenges to the objective of enabling access to and re-use of publicly funded research data as OER. Understanding this legal dimension and managing it effectively is essential in facilitating access to and re-use of research data.

A primary consideration is that much research data is subject to intellectual property rights which may be exercised by the rights owner to prevent or limit access and re-use.⁷ Where research data is copyright protected, rights to use or re-use it are subject to the scope of legitimate use prescribed by the law, which is usually restricted to particular users and specific acts or purposes (such as permitted under the fair dealing exception for research and study).⁸ Data users must obtain permission from data owners for access, use and re-use beyond the scope of the uses permitted under the legislation.⁹ Obtaining relevant permissions and negotiating licenses for research data can be a costly and time consuming process which may effectively preclude its use or re-use.¹⁰ If publicly funded research data are to be made available as OER, all intellectual property rights must be identified and managed in a manner consistent with, and that gives effect to open access and re-use objectives.¹¹

Legal challenges also arise from ambiguities about ownership of publicly funded research data. Researchers who generate data may be parties to contractual relationships with public funding agencies, other research collaborators, employers and research institutions.¹² In the context of a complex web of contractual relationships, researchers are often unsure about who owns the data or who has the right to authorise access to and re-use of it.¹³ There are also legal challenges arising from confidentiality, privacy and national security laws. Researchers are frequently subject to contractual, statutory or common law obligations to maintain the confidentiality of their research findings, including the data generated by the research project. Information privacy laws often present seemingly insurmountable barriers as access to and re-use of identified/identifiable personal information without the consent of the persons to whom the data relates will violate their right to informational privacy. Statutory protection of national security interests would prevent the disclosure of research data which would be prejudicial to national security.¹⁴

Novelty requirements in patent law are also relevant in determining how, and when, to provide access to research data. Premature disclosure of research data will preclude the patenting of an invention if the disclosed research data is considered as prior art and renders the invention no longer novel.¹⁵ Novelty requirements encourage researchers to restrict, limit, delay or withhold disclosure of the research data prior to the filing of a patent application.¹⁶ Lack of a legal duty to ensure data quality presents yet another legal challenge. The laws in most countries do not impose any duty of care on data providers who release the research data voluntarily, for free or without profit to the public.¹⁷ Therefore, data users need to consider the risks of obtaining incomplete, unfit, inaccurate or erroneous research data.¹⁸

3. ADDRESSING THE LEGAL CHALLENGES

To overcome the potential legal impediments to enabling access to and reuse of publicly funded data as OER, the following best practices are proposed for adoption:

- i. The relevant open access policy should make it clear that, subject to restrictions required for confidentiality, privacy and national security purposes, the owner/custodian of publicly funded data is required to permit access, use and re-use, including for data that is protected by intellectual property rights.¹⁹
- ii. Guidelines should clarify ownership of publicly funded research data created by: i) researchers in and outside the course of employment; ii) non-employee researchers; and iii) researchers involved in research collaborations with other researchers.
- iii. Publicly funded research data should be deposited in an open access repository following the expiry of an embargo period which allows data exclusivity. The duration of the embargo period depends on the requirements of the public research funding agency but, where it is not specified by the funding agency, data release should occur:
 - a. not later than two years from the collection/creation of the research data; or
 - b. immediately upon the first publication based on the research data; or
 - c. not later than one year from the end (either by expiry or termination) of the award/grant which funds the collection/creation of the research data; or
 - d. not later than one year upon completion of the research project for which the research data is collected/created.
- iv. Rights to use publicly funded research data should expressly permit acts that include the following:
 - a. quoting long excerpts;
 - b. distributing full-text copies to students and colleagues;
 - c. making copies on CDs for bandwidth-poor regions;
 - d. distributing semantically-tagged or otherwise enhanced (modified) versions;
 - e. migrating to new formats or media to ensure that documents remain readable as technologies change;
 - f. creating and archiving copies for long term preservation;
 - g. including works in a database or mash-up;
 - h. translation into another language; and
 - i. copying for the purposes of indexing, data mining and other kinds of processing.²⁰
- v. Publicly funded research data which are protected by copyright and which are released as OER should be licensed under a Creative Commons Licence (CC Licence). The most liberal CC Licence, Creative Commons Attribution (CC BY) - which reserves only the right to be attributed as data owner - should be adopted as the default licence for OER data.²¹
- vi. Publicly funded research data which contains confidential information should only be released where measures are in place to protect the confidentiality of the information. Such methods include data suppression, data random perturbations, and data coding and recoding. Where it is not appropriate or possible to use such methods to protect the confidential

- information, access to and reuse of publicly funded research data as OER must not be provided.
- vii. Publicly funded research data which contains direct/indirect identifiers or sensitive personal information of identified/identifiable subjects must only be released in a form that protects the informational privacy of the subjects of the research data. Access to and re-use of the research data is enabled by applying data redaction techniques such as anonymisation/de-identification, pseudonymisation, obfuscation, perturbation or data generalisation.
 - viii. Disclosure of research data containing information which is classified as prejudicial to national security is strictly prohibited. To avoid uncertainty, a classification of research data must be developed. Publicly funded research data which is not classified as restricted may be released for access and re-use.
 - ix. A timeframe for patent applications to be filed must be fixed to avoid any unnecessary delay in data release. A decision to patent must be made by the rights holder within six (6) months after formal notification of the invention. Where the decision is made not to patent the invention, the research data must be immediately released. Where the decision is to patent the invention, the patent application should be filed within six (6) months from the date the decision was made, unless it is shown that it is not possible due to the complexity of the patent to be filed.
 - x. A standard of care to ensure data quality, applicable to all data providers should be developed. The duty to ensure the quality of the research data is shared between a data creator (“primary data provider”), a data owner (if different from a data creator) and a repository/an archive/an enclave centre where the research data is deposited (A data owner and data repository/archive/enclave centres are known as “secondary data providers”). A primary data provider must supply the metadata which enables data users to assess the quality of the research data. A repository/archive/enclave centre must ensure that the primary data provider declares whether the research data is subject to peer review in-line with accepted best practice. Where the research data is not subject to peer-review, the primary data provider must warn the data users about the fact.

4. CONCLUSION

Where research data are produced with public funds, there are plausible moral and economic arguments supporting the view that it should be made freely available for access and re-use.²² However, legal challenges arise when providing for access to and re-use of publicly funded research data, based on intellectual property rights, confidentiality and information privacy laws, and the lack of a legal duty to ensure data quality. If these legal challenges are not understood, addressed and effectively managed, they will hinder the achievement of open access objectives. By adopting the best practices described in this paper, some of the most significant potential legal impediments can be overcome, enabling publicly funded research data to be made available for access and reuse as OER.

-
- 1 OECD Principles and Guidelines for Access to Research Data from Public Funding 2007.
 - 2 Committee on Ensuring the Utility and Integrity of Research Data in Digital Age, 'Ensuring the Integrity, Accessibility, and Stewardship of Research Data in the Digital Age' (National Academy of Sciences, National Academy of Engineering and Institute of Medicine, 2009), 22.
 - 3 John Houghton, Colin Steele and Peter Sheehan, 'Research Communication Costs in Australia: Emerging Opportunities and Benefits: A Report to the Department of Education, Science and Training' (Centre for Strategic Economic Studies Victoria University, Melbourne, 2006).
 - 4 Witold Kwasnicki, *Knowledge, Innovation and Economy: An Evolutionary Exploration* (Edward Elgar Publishing Ltd., Cheltenham, UK, 1996) xi;
 - 5 See Henry W Chesbrough, 'The Era of Open Innovation' (2003) (Spring) MIT Sloan Management Review 37; Terry Cutler, 'Innovation and Open Access to Public Sector Information' in Brian Fitzgerald (ed), *Legal Framework for e-Research: Realising the Potential* (Sydney University Press, Sydney, 2008) 29.
 - 6 See Jon-Arild Johannessen, Bjorn Olsen and Johan Olaisen, 'Aspects of Innovation Theory Based on Knowledge-Management' (1999) 19 International Journal of Information Management 121; Robinson Esalimba and William New, 'Spurring Local Innovation in Africa By Improving Access to Information' (2009), <<http://www.ip-watch.org/2009/10/19/spurring-local-innovation-in-africa-by-improving...>> (at 25 March 2010).
 - 7 Joan E Sieber, 'Social Scientists Concerns About Sharing Data' in Joan E Sieber (ed), *Sharing Social Science Data: Advantages and Challenges* (Sage Publications, Newbury Park, California, 1991) 143.
 - 8 See Ida Madieha Abdul Ghani Azmi 'Institutional Repositories in Malaysia: The Copyright Issues' (2009) 17(3) International Journal of Law and Information Technology 273; Thomas F Cotter, 'Transformative Use and Cognizable Harm' (2010) 12(4) Vanderbilt Journal of Entertainment and Technology Law 701.
 - 9 Linda Wang, 'Use of Images for Commercial Purposes: Copyright Issues Under Malaysian Laws' in Barbara Hoffman (ed), *Exploiting Images and Image Collections in the New Media: Gold Mine or Legal Minefield?* (Kluwer Law International, London, UK, 1999) 86.
 - 10 Peter Suber, *Open Access* (MIT Press, Cambridge, Massachusetts 2012) 68; Academic Senate of the California State University, 'Intellectual Property, Fair Use, and the Unbundling of Ownership Rights' (California State University, 2003) 18;
 - 11 Paul A David, 'Can 'Open Science' be Protected from the Evolving Regime of IPR Protections?' (2004) 129(March) Journal of Institutional and Theoretical Economics 1; Andrew Charlesworth, 'IPR and Research Data' (2011) Intellectual Property Rights and Research in the Digital Age, <<http://www.lib.cam.ac.uk/dataman/pages/IPR.html>> (at 28 June 2011).
 - 12 Ann Monotti and Sam Ricketson, *Universities and Intellectual Property: Ownership and Exploitation* (Oxford University Press, New York, 2003) 97.
 - 13 Kim Ji-Hyun, *Faculty Self-Archiving Behavior: Factors Affecting the Decision to Self-Archive* (PhD Dissertation Thesis, The University of Michigan, 2008) 54, 213.
 - 14 Stephen Hilgartner and Sherry I Brandt-Rauf, 'Data Access, Ownership, and Control: Toward Empirical Studies of Access Practices' (1994) 15 Science Communication 356; Alasdair Roberts, 'National Security and Open Government' (2004) 9(2) Georgetown Public Policy Review 69.
 - 15 Margo A Bagley, 'Academic Discourse and Proprietary Rights: Putting Patents in Their Proper Place' (2006) XLVII(2) Boston College Law Review 218.
 - 16 Nathaniel B Lipkus, Jocelyn E Mackie and Peter A Singer, 'Guidance for Reconciling Patent Rights and Disclosure of Findings at Scientific Meetings' (2010) 8(15) Health Research Policy and Systems 4.
 - 17 Cheryl Foong, 'Open Content Licensing of Public Sector Information and the Risk of Tortious Liability for Australian Governments' (2010) 17(2) Murdoch University Electronic Journal of Law 23, <<https://elaw.murdoch.edu.au/index.php/elawmurdoch/issue/current>> (at 17 August 2011).
 - 18 Raka Banerjee, 'Open Data is Not Enough' (2011), <<http://blogs.worldbank.org/developmenttalk/open-data-is-not-enough-0>> (at 16 August 2011).
 - 19 See James Manyika et al, 'Big Data: The Next Frontier for Innovation, Competition, and Productivity' (McKinsey Global Institute, 2011) 120; Brian Rappert and Andrew Webster,

-
- 'Regimes of Ordering: The Commercialization of Intellectual Property in Industrial-Academic Collaborations' (1997) 9(2) *Technology Analysis & Strategic Management* 116.
- 20 Suber, above n 10, 73.
- 21 See Graham Greenleaf, 'Unlocking IP to Stimulate Australian Innovation: An Issues Paper' (University of New South Wales, 2008); Paul Uhlir and Peter Schroder, 'Open Data for Global Science' in Brian Fitzgerald (ed), *Legal Framework for e-Research: Realising the Potential* (Sydney University Press Sydney, 2008) 209; Brian Fitzgerald and Kylie Pappalardo, 'The Law as Cyberinfrastructure' (2007) 3(3) *CT Watch Quarterly* 51.
- 22 See Wijayananda Jayaweera, 'Whose Knowledge?' (2001) 28(1) *Media Asia* 22; Rick Anderson, 'Author Disincentives and Open Access' (2004) 30(4) *Serials Review* 288.

REFERENCES

- Abdul Ghani Azmi, IM, 'Institutional Repositories in Malaysia: The Copyright Issues' (2009) 17(3) *International Journal of Law and Information Technology* 273
- Academic Senate of the California State University, 'Intellectual Property, Fair Use, and the Unbundling of Ownership Rights' (California State University, 2003) 18
- Anderson, R, 'Author Disincentives and Open Access' (2004) 30(4) *Serials Review* 288
- Bagley, MA, 'Academic Discourse and Proprietary Rights: Putting Patents in Their Proper Place' (2006) XLVII(2) *Boston College Law Review* 218
- Banerjee, R, 'Open Data is Not Enough' (2011), <<http://blogs.worldbank.org/developmenttalk/open-data-is-not-enough-0>> (at 16 August 2011)
- Charlesworth, A, 'IPR and Research Data' (2011) *Intellectual Property Rights and Research in the Digital Age*, <<http://www.lib.cam.ac.uk/dataman/pages/IPR.html>> (at 28 June 2011)
- Chesbrough, HW, 'The Era of Open Innovation' (2003) (Spring) *MIT Sloan Management Review* 37
- Committee on Ensuring the Utility and Integrity of Research Data in Digital Age, 'Ensuring the Integrity, Accessibility, and Stewardship of Research Data in the Digital Age' (National Academy of Sciences, National Academy of Engineering and Institute of Medicine, 2009)
- Cotter, TF, 'Transformative Use and Cognizable Harm' (2010) 12(4) *Vanderbilt Journal of Entertainment and Technology Law* 701
- Cutler, T, 'Innovation and Open Access to Public Sector Information' in Brian Fitzgerald (ed), *Legal Framework for e-Research: Realising the Potential* (Sydney University Press, Sydney, 2008)
- David, PA, 'Can 'Open Science' be Protected from the Evolving Regime of IPR Protections?' (2004) 129(March) *Journal of Institutional and Theoretical Economics* 1
- Esalimba, R, New, W, 'Spurring Local Innovation in Africa By Improving Access to Information' (2009), <<http://www.ip-watch.org/2009/10/19/spurring-local-innovation-in-africa-by-improving...>> (at 25 March 2010)
- Fitzgerald, B, Pappalardo, K, 'The Law as Cyberinfrastructure' (2007) 3(3) *CT Watch Quarterly* 51
- Foong, C, 'Open Content Licensing of Public Sector Information and the Risk of Tortious Liability for Australian Governments' (2010) 17(2) *Murdoch University Electronic Journal of Law* 23, <<https://elaw.murdoch.edu.au/index.php/elawmurdoch/issue/current>> (at 17 August 2011)

-
- Greenleaf, G, 'Unlocking IP to Stimulate Australian Innovation: An Issues Paper' (University of New South Wales, 2008)
- Hilgartner, S, Brandt-Rauf, SI, 'Data Access, Ownership, and Control: Toward Empirical Studies of Access Practices' (1994) 15 *Science Communication* 356
- Houghton, J, Steele, C, Sheehan, P, 'Research Communication Costs in Australia: Emerging Opportunities and Benefits: A Report to the Department of Education, Science and Training' (Centre for Strategic Economic Studies Victoria University, Melbourne, 2006)
- Jayaweera, W, 'Whose Knowledge?' (2001) 28(1) *Media Asia* 22
- Johannessen, J-A, Olsen, B, Olaisen, J, 'Aspects of Innovation Theory Based on Knowledge-Management' (1999) 19 *International Journal of Information Management* 121
- Kim, JH, Faculty Self-Archiving Behavior: Factors Affecting the Decision to Self-Archive (PhD Dissertation Thesis, The University of Michigan, 2008)
- Kwasnicki, W, *Knowledge, Innovation and Economy: An Evolutionary Exploration* (Edward Elgar Publishing Ltd., Cheltenham, UK, 1996)
- Lipkus, NB, Mackie, JE, Singer, PA, 'Guidance for Reconciling Patent Rights and Disclosure of Findings at Scientific Meetings' (2010) 8(15) *Health Research Policy and Systems* 4
- Manyika, J, Chui, M, Brown, B, Bughin, J, Dobbs, R, Roxburgh, C, Byers, AH, 'Big Data: The Next Frontier for Innovation, Competition, and Productivity' (McKinsey Global Institute, 2011)
- Monotti, A, Ricketson, S, *Universities and Intellectual Property: Ownership and Exploitation* (Oxford University Press, New York, 2003)
- OECD Principles and Guidelines for Access to Research Data from Public Funding 2007
- Rappert, B, Webster, A, 'Regimes of Ordering: The Commercialization of Intellectual Property in Industrial-Academic Collaborations' (1997) 9(2) *Technology Analysis & Strategic Management* 116.
- Roberts, A, 'National Security and Open Government' (2004) 9(2) *Georgetown Public Policy Review* 69
- Sieber, JE, 'Social Scientists Concerns About Sharing Data' in Joan E Sieber (ed), *Sharing Social Science Data: Advantages and Challenges* (Sage Publications, Newbury Park, California, 1991)
- Suber, P, *Open Access* (MIT Press, Cambridge, Massachusetts 2012) 68
- Uhlir, P, Schroder, P, 'Open Data for Global Science' in Brian Fitzgerald (ed), *Legal Framework for e-Research: Realising the Potential* (Sydney University Press Sydney, 2008) 209
- Wang, L, 'Use of Images for Commercial Purposes: Copyright Issues Under Malaysian Laws' in Barbara Hoffman (ed), *Exploiting Images and Image Collections in the New Media: Gold Mine or Legal Minefield?* (Kluwer Law International, London, UK, 1999) 86