

This paper was presented at the conference:

Oxford eResearch 2008

11-13 September 2008

University of Oxford

Conference website:

<http://www.oii.ox.ac.uk/microsites/ereseach08/index.cfm>

Conference papers collection:

<http://ora.ouls.ox.ac.uk/objects/uuid%3A64aa6f39-7e81-4d42-a008-ee2d7524bd67>

Conference organisers:

Oxford Internet Institute

University of Oxford

1 St Giles, Oxford OX1 3JS

<http://www.oii.ox.ac.uk/>

email: enquiries@oii.ox.ac.uk

Oxford e-Research Centre

University of Oxford

7 Keble Road, Oxford OX1 3QG

<http://www.oerc.ox.ac.uk/>



LEGAL FRAMEWORK FOR e-RESEARCH: REALISING THE POTENTIAL

**EDITED BY
BRIAN FITZGERALD**

BA (Griff), LLB (Hons) (QUT), BCL (Oxon), LLM (Harv), PhD (Griff)
Professor of Intellectual Property Law and Innovation (QUT Faculty of
Law), Project Leader of the Legal Framework for e-Research Project
Barrister, Supreme Court of Queensland and High Court of Australia



SYDNEY UNIVERSITY PRESS

Published by
Sydney University Press
University of Sydney
NSW, 2006 Australia
www.sup.usyd.edu.au

Publication date: August 2008

© Copyright individual authors, 2008

© Sydney University Press 2008

The material in this publication is based on papers presented at the Legal Framework for e-Research International Congress convened by the Queensland University of Technology (<http://www.qut.edu.au>) at the Gold Coast, Australia, July 2007. This publication is an output of the Australian Government Department of Education, Employment and Workplace Relations (DEEWR) funded Legal Framework for e-Research Project led by Professor Brian Fitzgerald and based at the Queensland University of Technology.

Licence information

This work (with the exception of Chapter 2: The Fifth Dimension and Chapter 9: NIH Data and Resource Sharing, Data Release and Intellectual Property Policies for Genomics Community Resource Projects) is licensed under an Australian Creative Commons 'By Attribution', 'Non Commercial', 'No Derivative Works.' 2.5 Licence
<<http://creativecommons.org/licenses/by-nc-nd/2.5/au/>>.

See <http://creativecommons.org.au/licenses> for more details.

Chapter 2: The Fifth Dimension Licence Information

Where copyright subsists in this chapter, which is a United States Government Work under s 105 of the U.S. Copyright Act, the work is licensed under a United States Creative Commons 'By Attribution' 3.0 License <<http://creativecommons.org/licenses/by/3.0/us/>>.

Chapter 9: NIH Data and Resource Sharing, Data Release and Intellectual Property Policies for Genomics Community Resource Projects Licence Information

Where copyright subsists in this chapter, which is a United States Government Work under s 105 of the U.S. Copyright Act, the work is licensed under a United States Creative Commons 'By Attribution' 3.0 License <<http://creativecommons.org/licenses/by/3.0/us/>>.

ISBN 978-1920898-93-9 pbk

ISBN 978-1-920899-16-5 hbk

The digital version of this book is also available electronically through the Sydney eScholarship Repository (<http://ses.library.usyd.edu.au>) and

The Queensland University of Technology ePrints Repository at:
(<http://eprints.qut.edu.au>).

PREFACE

This book is the product of a conference that I convened at the Gold Coast Australia on the 11th and 12th of July 2007, titled *Legal Framework for e-Research*.¹

The conference was undertaken as part of the research program of the Legal Framework for e-Research Project. The conference, the project and this book have been made possible with the support and sponsorship of the federal Department of Education Science and Training (DEST) which since 2008 has been restructured into the new departments of Education, Employment and Workplace Relations (DEEWR) and Innovation, Industry, Science and Research (DIISR).

In the process of running the conference and preparing this book for publication I am thankful for the tremendous support of Dr Evan Arthur, Margot Bell and Clare McLaughlin from DEST, Professor Tom Cochrane DVC QUT, Dr Terry Cutler (Cutler & Co.), Dr Mary O’Kane (O’Kane & Associates), Dr Fiona Stanley (Telethon Institute for Child Health), Dr Chris Greer (NSF), Claire Driscoll (NIH), Dr Richard Jefferson (CAMBIA), Dr Michael Spence Vice Chancellor University of Sydney, Professor John Unsworth (UIUC), Fred Friend (JISC), Steve Matheson (ABS), Paul Uhler (National Academies) John Wilbanks (Science Commons), Professor Mark Perry (UWO), Dr Amanda McBratney and Malcolm McBratney (McCullough Robertson Lawyers), James Casey (UIDP), Adam Liberman (CSIRO), Andrew Hayne (OPC), David Ruschena (Health Legal), Professor Bernard Pailthorpe (UQ), Philip Crisp (AGS), Maree Heffernan (Qld Government), Professor Paul David (Stanford), the distinguished speakers and participants who attended the conference, my team at QUT, including Professor Anne Fitzgerald, Scott Kiel-Chisholm, Nic Suzor, Damien O’Brien, Kylie Pappalardo, Tanya Butkovsky, Amy Barker, Dr Annie Connell and Anthony Austin and the people at Sydney University Press especially Susan Murray-Smith and Ross Coleman.

¹ For the full program see: <<http://www.eresearch.law.qut.edu.au/conference>>

In particular I would like to acknowledge the work of my research assistant on this project, Celeste Bennett, who has done an excellent job on the chapters in this book.

The Gold Coast conference was followed by the *Australian National Summit on Open Access to Public Sector Information*² held at the Queensland Parliament House on the 13th of July 2007. Dr Cutler's paper in this book is drawn from that event.

This book aims to document the considerable insights and ideas that were offered at the Gold Coast conference. The majority of the papers in the book have been prepared as full academic papers (a number already published elsewhere) while others are shorter conference style papers with one being derived from a transcript of the proceedings.

The Legal Framework for e-Research project is based at QUT and aims to enhance the enormous potential of e-Research by ensuring law can work as an enabler in this fast paced, networked and serendipitous environment. It grew out of a CODATA conference I attended at UNESCO in Paris in 2004 – International Workshop on the Information Commons for e-Science³ – and my realisation that we were not doing enough in this area in Australia.

The project is due for completion in late 2008 and can only be seen as a precursor to the further work that will be done in this area. A national survey of attitudes and practices, the Gold Coast conference, a report on the legal framework for data management, this book and a high level international roundtable on negotiating and concluding collaborative research agreements held at QUT in June 2008 are major outputs of this project so far.⁴

² For the full program see:

[http://datasmart.oesr.qld.gov.au/Events/datasmart.nsf/0/D2D95818B020049F4A25732C0006EE9A/\\$FILE/Australian%20National%20Summit%20On%20Open%20Access%20Program.pdf?openelement](http://datasmart.oesr.qld.gov.au/Events/datasmart.nsf/0/D2D95818B020049F4A25732C0006EE9A/$FILE/Australian%20National%20Summit%20On%20Open%20Access%20Program.pdf?openelement) For the Conference Report see:

<[http://www.qsic.qld.gov.au/QSIC/QSIC.nsf/0/D6C8E0616BC7FB414A2573B7000C42E5/\\$FILE/Conference%20Report%20%20National%20Summit%20Open%20Access.pdf?openelement](http://www.qsic.qld.gov.au/QSIC/QSIC.nsf/0/D6C8E0616BC7FB414A2573B7000C42E5/$FILE/Conference%20Report%20%20National%20Summit%20Open%20Access.pdf?openelement)>

³ See generally: <http://portal.unesco.org/es/ev.php-URL_ID=28723&URL_DO=DO_TOPIC&URL_SECTION=201.html>

⁴ See generally: <<http://www.eresearch.law.qut.edu.au>>

I hope you enjoy reading the material in this book and that it provides you with a framework for understanding key legal issues in the e-Research landscape.

Professor Brian Fitzgerald

Professor of Intellectual Property and Innovation QUT Law Faculty

Project Leader Legal Framework for e-Research Project

www.eresearch.law.qut.edu.au

Brisbane, June 30 2008

FOREWORD

The law famously trails behind the social and economic changes in human society which it is supposed to regulate. In a time of such rapid change as we have in the early 21st Century, efforts to reduce some of the undesirable outcomes of this lagging characteristic are to be commended, most particularly when those efforts are directed at improvements which are designed to support greater access to knowledge and the sharing of information in and among communities with increasingly complex problems requiring ever more urgent solution.

It is a commonplace to claim that intellectual property laws are arcane. Nowhere is this more ably demonstrated than in the convoluted attempts to reform copyright law across a range of jurisdictions in the last decade. Australia is no small example, with amendments to the *Copyright Act 1968* in 2000 designed to address the requirements of the digital environment, followed by a range of other changes including the significant realignment of Australian copyright law with that of the domestic United States' legal environment brought about as a result of the *Australia-United States Free Trade Agreement* (2004).

The failure of the law to keep pace with the real and dynamic changes in the virtual environments in which our educators, researchers, and citizens generally find themselves is understandable, as new ways of doing things arrive so quickly that it takes time for society at large to adequately comprehend and understand them. And, no sooner than one set of conditions has been understood, there arrive new and sometimes contradictory developments which demand a rethink of any regulatory creativity that might be based on the previous set of conditions.

But having said this, it is also the case that significant coalitions of interests are able to form a general direction to follow in the advocacy that they pursue with law makers.

One such coalition is that which can be generally described by research and education communities, worldwide. Contradictions and conundrums that have developed in the digital age over access to research outputs (for example, the idea of the expensive, subscription bundle of electronic journals), have met with countervailing influences,

policies and developments which are aimed at exploiting new possibilities for the sharing of knowledge. With this general aim, the Open Access movement has developed, which at its simplest merely urges researchers who are not publishing for individual monetary return to release the quality-certified (i.e. peer reviewed) outputs of their research online at the same time as they are submitted for publication in the toll gated journal literature.

At the same time as that movement for greater access has developed, so have we seen innovation in the area of intellectual property and copyright law. This includes the advent of open content licensing, which seeks to address the issue of confusion and limitation surrounding the understand of copyright requirements, by providing the creators of intellectual property with options to express in a positive way rights to use which are otherwise exclusively reserved for the creator. While there is debate about the extended intersection between open content licensing, and open access more generally, both movements, which are global in dimension, have the same end – the greater sharing of knowledge.

The Australian Government through the former Department of Education, Science and Training (DEST) has for some years been pursuing investments in Systemic Information Infrastructure (SII) to support in various ways research and access to research in Australia. Informing this has been the evolution of an Accessibility Framework which has been developed to guide the Government's intention to the see that taxpayer investment in research activities in our institutions is matched by appropriately available public access to the outcomes and outputs of such research.

While these developments have been occurring, the landscape generally described by the term “e-Research”, has also been rapidly evolving. The working definition of e-Research adopted by the Australian Government as it went about its Systemic Infrastructure investment program, and later its National Collaborative Research Infrastructure Strategy (NCRIS) investments, was to define it as encapsulating research activities that use “a spectrum of advanced ICT capabilities” and embrace(s) new research methodologies emerging from increased access to ... broadband ... networks, research instruments, ... software and infrastructure services

that enable secure connectivity and application tools that encompass discipline specific tools and interaction tools”.⁵

Under the Platforms for Collaboration portion of NCRIS, these investments have been further refined into significant heads of activity which have led to the development of a National Computation Initiative, an Australian Research Collaboration Service, and an Australian National Data Service.

At the same time, one of the investments made by the Commonwealth Government in advancing activity under the earlier banner of Systemic Information Infrastructure, was to support an initiative centred at the Queensland University of Technology, known as Open Access to Knowledge, (OAK) which led to further research and a new project – The Legal Framework for e-Research Project – on the challenges and problems that could be envisaged as e-Research activity gathers momentum, not only regionally and nationally, but worldwide.

The recognition that increasingly sophisticated ICT techniques would need to be matched by increasingly sophisticated ways of thinking about some of the legal issues that might arise, particularly given the minefield posed just by conditions relating to the collection and availability of data across a wide number of fields, therefore led to the initiative represented by the contents of this report. We need skilful thinking about the legal context for all of the rapid and new, indeed scarcely definable activities that are mushrooming in this sphere. An early start on this was the July 2007 conference on the topic.

It is my pleasure to commend the papers in this volume for those interested in considering the many interesting challenges posed for researchers, legal specialists and administrators in this area.

Professor Tom Cochrane

Deputy Vice Chancellor of QUT

Chair of the Australian e-Research Infrastructure Council (AeRIC)

Brisbane, 30 June 2008

⁵ Final Report of the e-Research Coordinating Committee, *An Australian e-Research Strategy and Implementation Framework* (2006),

TABLE OF CONTENTS

PREFACE.....	iii
FOREWORD.....	vi
Chapter 1 Introduction	1
<i>Professor Brian Fitzgerald</i>	
Part One – The Dimensions and Possibilities of e-Research	
Chapter 2 The Fifth Dimension	5
<i>Dr Chris Greer</i>	
Chapter 3 Innovation and Open Access to Public Sector Information	25
<i>Dr Terry Cutler</i>	
Chapter 4 Cyber Infrastructure for the Humanities and Social Sciences	40
<i>Professor John Unsworth</i>	
Part Two – Institutional and Legal Frameworks	
Chapter 5 Designing Institutional Infrastructures for e-Science	55
<i>Professor Paul David and Dr Michael Spence</i>	
Chapter 6 The Law as Cyber Infrastructure	147
<i>Professor Brian Fitzgerald and Kylie Pappalardo</i>	
Part Three – Data Ownership, Access and Reuse	
Chapter 7 Understanding the Legal Implications of Data Sharing, Access and Reuse in the Australian Research Landscape	161
<i>Professor Anne Fitzgerald, Kylie Pappalardo and Anthony Austin</i>	
Chapter 8 Open Data for Global Science	189
<i>Paul F Uhlir and Peter Schroder</i>	

Chapter 9 NIH Data and Resource Sharing, Data Release and Intellectual Property Policies for Genomics Community Resource Projects.....	223
<i>Claire Driscoll</i>	
Chapter 10 Cyberinfrastructure For Knowledge Sharing.....	240
<i>John Wilbanks</i>	

Part Four – Scholarly Communications

Chapter 11 Working for a Research-Friendly IPR Framework in the UK.....	257
<i>Fred Friend</i>	
Chapter 12 Creating a Legal Framework for Copyright Management of Open Access within the Australian Academic and Research Sector.....	264
<i>Professor Brian Fitzgerald, Professor Anne Fitzgerald, Professor Mark Perry, Scott Kiel-Chisholm, Erin Driscoll, Dilan Thampapillai and Jessica Coates</i>	

Part Five – Contractual and Technological Frameworks

Chapter 13 The University-Industry Demonstration Partnership: An Incremental Improvement to University-Industry Collaboration.....	333
<i>James Casey</i>	
Chapter 14 Streamlining Collaborative Agreements in an e-Research World.....	350
<i>Anthony Austin and Professor Brian Fitzgerald</i>	

Part Six – Privacy and e-Research

Chapter 15 A Win:Win for Data Access: Balancing Public Good with Privacy Concerns	399
<i>Professor Fiona Stanley AC</i>	
Chapter 16 Privacy Regulation and e-Research.....	410
<i>Andrew Hayne</i>	
Chapter 17 A Primer in the Politics of Privacy and Research	423
<i>David Ruschena</i>	

Part Seven – Patents and Commercialisation in an e-Research world: New Models

Chapter 18 Science as Social Enterprise: The CAMBIA BIOS

Initiative 437

Dr Richard Jefferson

Part Eight – International Dimensions

Chapter 19 e-Research and Jurisdiction 481

Gaye Middleton

Part Nine – Survey of Stakeholder Attitudes

Chapter 20 Australian Survey on Legal Issues Facing e-Research..... 497

Maree Heffernan and Scott Kiel-Chisholm

BIOGRAPHIES 525

INDEX..... 544

INTRODUCTION

Professor Brian Fitzgerald¹

The conference and research project that produced this book have operated throughout 2006, 2007 and 2008 in a rapidly changing knowledge landscape. One of the most significant changes is that e-Research has moved from being a specialist activity or technique to one that now promises to be adopted as a methodology for almost all research.

Over the last three years we have seen an enormous uptake in the use of online social networking tools and other Web 2.0 modalities. Innovation policy around the world has moved to embrace the idea that the exchange of ideas through networks is now a critical factor to finding new ways of doing things and enhancing productivity.² Notions such as “user generated”, “peer produced” and “collaboration” or “crowd sourcing” which are hallmarks of the Web 2.0 era are rapidly becoming part of mainstream research methods.

Apparent in this changing landscape is the need to be able to foster a pure, dynamic and serendipitous exchange of ideas – information flows. For the most part the technology can provide this opportunity yet the law remains as a potential roadblock. In response to this challenge this book aims to explore how the law might flow with the technology to promote and sponsor – enable rather than inhibit – the dynamic and enormous potential of e-Research.

¹ Professor of Intellectual Property and Innovation (QUT Law Faculty); Project Leader Legal Framework for e-Research (QUT)

² Sacha Wunsch-Vincent and Graham Vickery, *Participative Web and User-Created Content: Web 2.0, Wikis and Social Networking* (2007) Organisation for Economic Co-operation and Development, available at <http://www.oecd.org/dataoecd/57/14/38393115.pdf>; OECD, *Shaping Policies for the Future of the Internet Economy*, Annexes, available at <http://www.oecd.org/dataoecd/1/28/40821729.pdf>; OECD, *Seoul Declaration on the Future of the Internet Economy*, 18 June 2008, available at <http://www.oecd.org/dataoecd/49/28/40839436.pdf>

Part One of this book considers the broader prospect and context of what e-Research will allow. Part Two looks more closely at the role law will play in the e-Research environment. Part Three focuses on the key issues of data exchange and data management highlighting important legal issues. Part Four reflects on the changing nature of Scholarly Communications while Part Five looks at the fundamental role of agreements for collaborative endeavour (contracts) in structuring collaboration and calls for greater consideration of way we can streamline the process. Part Six examines the role and operation of privacy law in an e-Research world while Part Seven posits a new approach to commercialisation that embraces the paradigm of open innovation. Part Eight looks at the international legal implications for e-Research and Part Nine considers the national survey we undertook on e-Research, collaborative agreements and data management.

Some of the key themes that emerge in this book are that:

- e-Research provides tremendous opportunity yet to realise its full potential we need to support it through forward looking institutional, legal and policy frameworks
- Data exchange is a critical yet potentially overwhelming exercise in this new environment so we must work hard to facilitate data management, interoperability, access and reuse
- Privacy, intellectual property and agreement issues present immediate yet not insurmountable challenges. They need to be considered in terms of potential law reform, institutional and policy development and the practices employed in the research sector
- e-Research is a global activity so our approaches and solutions must be cognisant of global trends and practices in law and policy
- Ultimately research is originated by people and in building the e-Research platform we are embedded in a network of human as well as technological relationships. We need to understand how these relationships exist and how they might be nurtured and accommodated or might even change in this new landscape.

The underlying theme of this book, as evidenced in the chapters that follow, is that law is part of the infrastructure in the e-Research environment in that it provides pathways for activity to occur – what we might call “law as cyberinfrastructure”. In looking at the law in this way we highlight the positive role law might play in sponsoring or enabling innovation in the e-Research world

PART ONE

THE DIMENSIONS AND
POSSIBILITIES OF e-RESEARCH

THE FIFTH DIMENSION

Chris Greer¹

INTRODUCTION

The aim of this chapter is to consider a five-dimensional world made possible by cyberinfrastructure and how this notion influences legal frameworks. In discussing this five-dimensional world I will highlight fundamental challenges that hinder this vision, which is a shared vision, not unique to the National Science Foundation, but common to countries throughout the world. I will also consider strategies that could assist in achieving this fifth dimension.

Definitions

Throughout this chapter I use the term ‘cyberinfrastructure’. Fran Berman defines cyberinfrastructure as the ‘coordinated aggregate of software, hardware and other technologies, as well as human expertise, required to support current and future discoveries in science and engineering.’² This definition is particularly appropriate in the context of the fifth-dimension because the definition encompasses not just hardware, software and network fabric but organisations, people and

Thank you very much for inviting me to present at *The Legal Framework for e-Research Conference* at the Queensland University of Technology. It has provided me with an opportunity to interact with people who are making significant contributions in the area of data preservation and legal frameworks for information integrations. I have greatly appreciated this opportunity and to hear from you and share what the National Science Foundation is hoping to achieve.

This chapter is derived from a transcript of a presentation given by Dr Chris Greer at the *Legal Framework for e-Research* conference convened by the Queensland University of Technology Law Faculty in 2007.

¹ Senior Advisor for Digital Data in the Office of Cyberinfrastructure, National Science Foundation (NSF).

² Fran Berman, ‘Workshop Concept’ (SBE/CISE Workshop on Cyberinfrastructure for the Social Sciences) <http://vis.sdsc.edu/sbe/SBE-CISE_Workshop_Intro.pdf>.

their expertise which makes all of this possible, and may be considered the most integral part of cyberinfrastructure.

I also use the term ‘data’. Data refers to items that can be digitised, stored in digital form and accessed electronically. This includes numeric information or text, as well as images, audio, algorithms, software, simulations to name a few.

HOW IS CYBERINFRASTRUCTURE CHANGING OUR LIVES?

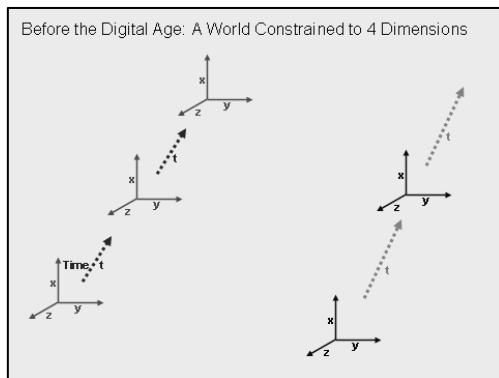
The National Science Foundation believes that:

The conduct of science and engineering is changing and evolving. This is due, in large part, to the expansion of networked cyberinfrastructure . . .³

The fundamental question that should be asked is in what ways are science and engineering changing, and what are the driving forces for those changes?

Prior to the digital age people operated in a world constrained by four dimensions, particularly the three dimensions of place and one dimension of time.

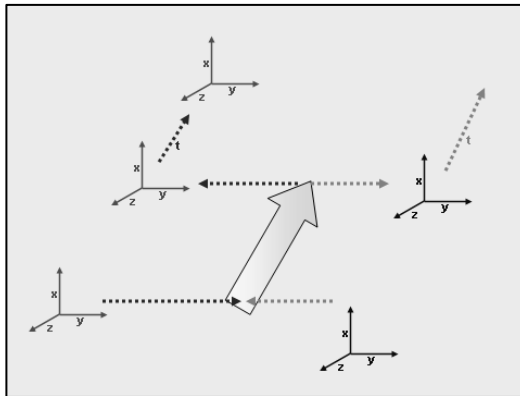
Figure One



³ National Science Foundation, *National Science Foundation Strategic Plan 2006–2011*.

Figure One has two trajectories, one of which may have been your trajectory as you prepared for a meeting, the other trajectory may have been the trajectory of the person you were meeting. For both of you to meet you had to agree to interrupt your trajectories, stipulate a time to meet, and then return to your separate trajectories at the end of the meeting (see Figure 2). That is the world that people are accustomed to operating in.

Figure Two



Cyberinfrastructure creates a fifth dimension that is present alongside the existing four dimensions. This fifth dimension provides people with the opportunity to search for information they did not know existed, in places they will never visit, while interacting with unknown people in other places, using instruments they do not own and do not know how to operate in a deep, technical level, but which they have access to because of cyberinfrastructure.

This fifth dimension also allows people to meet in a synchronised mode at an agreed time, or in a meta-synchronised mode where people who are operating at a distance, in their own time zones and context can interact with one another (see Figure Three).

Figure Three

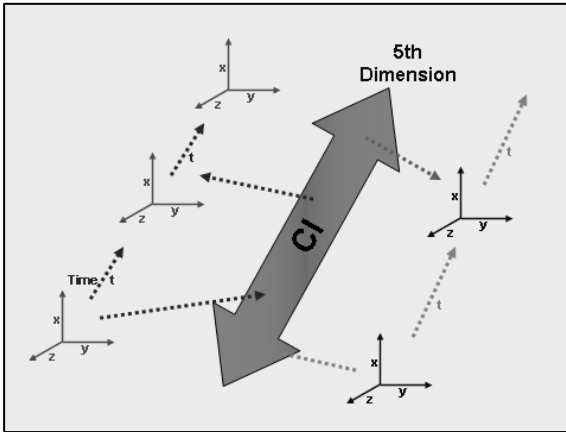
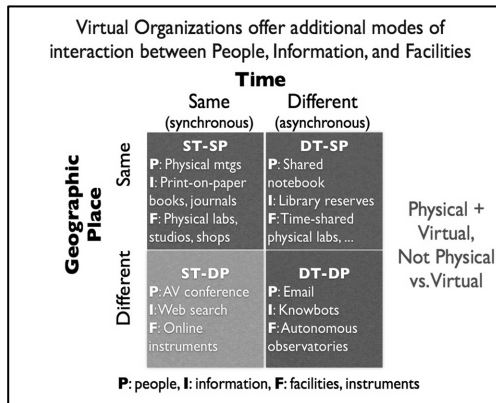


Figure Four is a two-by-two matrix. In a four dimensional world people spend the majority of their time operating in the same-place same-time sector. However other sectors, particularly the different-time different-place sector, have become available for activities involving people-information-facilities and this provides for opportunities that do not exist in the same-place same-time mode. This expansion into other areas of the matrix that is occurring is what is meant by ‘operating in a world of five dimensions’.

Opening a fifth dimension through cyberinfrastructure is the defining feature of the digital age. Most people have read Thomas Friedman’s *The World is Flat*⁴ and are familiar with the idea of dialling up a helpdesk and speaking with someone in India or Malaysia. Software development activities involve teams scattered around the globe so that the development cycle moves with daylight around the globe to become a 24-hour, seven-day a week activity. These concepts are what Thomas Friedman was referring to when he described the world as flat. However this is only part of the picture, because this flat world is also expanding.

⁴ Thomas Friedman, *The World is Flat: A Brief History of the Twenty-First Century* (2005).

Figure Four



Source: Dr. Daniel E. Atkins

Prior to the advent of network cyberinfrastructure there was no economic space for companies such as Google and Amazon, nor was there a place for the National Virtual Observatory⁵ which is vital to astronomy. These new spaces and opportunities, whether economic, scientific or educational, arise because of cyberinfrastructure: the world is getting bigger.

Unlike the four dimensional world in which the driving forces for progress were physical and economic assets, the five dimensional world's primary drivers for progress are information assets and the critical driver for progress in this world is the ability to use information in integrative and innovative ways.

FUNDAMENTAL CHALLENGES OF OUR TIME

Some of the major challenges of our time are enabled by cyberinfrastructure for information integration, such as the ability to ask and answer questions like how and where did life arise on earth? This answer can only be answered using a combination of scientific areas

⁵ United States National Virtual Observatory <<http://www.us-vo.org/>>.

such as systematic biology, palaeobiology, biochemistry, metabolic biochemistry, genomics and geochemistry. Likewise the question, what is the biological basis of consciousness? This answer will require integrating information from biology and other sciences.

Some of the large questions of our time will require using information from a wide variety of sources and frameworks together in an integrative way. It is this ability, the ability to integrate information, which is critical to being successful in a five-dimensional world. Individuals, groups and nations that fail to fully embrace this five dimensional world will fall behind.

CHARACTERISTICS OF A FIVE-DIMENSIONAL WORLD

Some of the primary characteristics of a five-dimensional world include: barriers of time and place, which are characteristic of a four-dimensional world, are reduced, information is a primary driver for progress, access to information is available to specialists and non-specialists alike and that the realm of the possible is expanded through new capabilities, resources and mechanisms.

At the National Science Foundation, as in other science and education, engineering and research organisations, the increase of digital outputs is on the rise.

An example of this is from astronomy, where it has becoming increasingly apparent that the dynamics of the universe are an important element for study. The Large Synoptic Survey Telescope is a project that is expected to be realised sometime in the 2010s.⁶ The telescope will map the sky every night using a three billion pixel camera, taking a full survey of the sky in under a week.⁷ On a clear night this effort will generate 30 terabytes of data.⁸

⁶ 'Steering the Future of Computing' (23 March 2006)440 *Nature* 383
<<http://www.nature.com/nature/journal/v440/n7083/pdf/440383a.pdf>>.

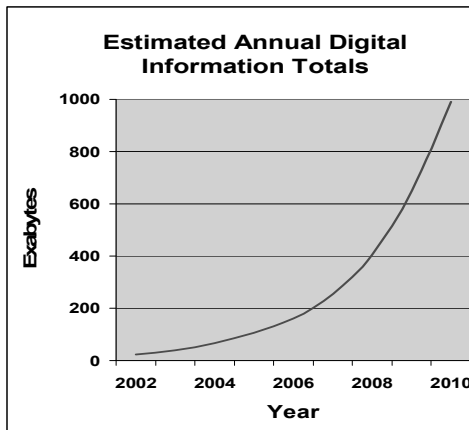
⁷ 'Steering the Future of Computing' (23 March 2006)440 *Nature* 383
<<http://www.nature.com/nature/journal/v440/n7083/pdf/440383a.pdf>>; Large Synoptic Survey Telescope <http://www.lsst.org/lsst_home.shtml>.

⁸ 'Steering the Future of Computing' (23 March 2006)440 *Nature* 383
<<http://www.nature.com/nature/journal/v440/n7083/pdf/440383a.pdf>>.

In biology the National Ecological Observatory Network,⁹ and in the geosciences and the climate sciences the Global Earth Observation System of Systems¹⁰ are examples of efforts generating datasets analogous to this sort of magnitude. This is also true for society as a whole.

Figure Five¹¹ is a projection from an International Data Corporation study, which extrapolates the previous work of Michael Laskey of Rutgers University and Hal Varian and Peter Lyman at the University of California at Berkley. The projection shows in exabytes by a year the amount of digital information that is generated globally.

Figure Five



In 2006 a total of 161 exabytes were generated around the world, this is more information than all the documents in the previous 40 000 years human history contained. Printed out in volumes of ones and zeros this would equal a stack of 12 volumes reaching from the surface of the earth to the surface of the sun.

⁹ National Ecological Observatory Network <<http://www.neoninc.org/>>.

¹⁰ Global Earth Observation System of Systems <<http://www.epa.gov/geoss/>>.

¹¹ International Data Corporation, 'The Expanding Digital Universe' (IDC White Paper sponsored by EMC Corporation, March 2007) <<http://www.emc.com/collateral/analyst-reports/expanding-digital-idc-white-paper.pdf>>; Peter Lyman and Hal R Varian, *How Much Information* (2003) <<http://www.sims.berkeley.edu/how-much-info-2003>>.

The curve in Figure Five is exponential, which indicates that linear solutions, such as expert curation models, will not be adequate in addressing this growth problem. The curve also illustrates that given the volume of information the vast majority will never be seen by human eyes. The information will be passed, sorted, filtered, analysed and reduced to a level humans can understand. Volume is an important challenge in the world of five dimensions and predicts the need for exponential solutions.

Figure Six¹² is a summary of information technologies over the course of human history from stone, clay and papyrus to paper and now digital forms of storage.

Figure Six

Diagram omitted from digital version

Please see hard copy for diagram. For hard copy, please contact Sydney University Press (www.sup.usyd.edu.au)

Source: Berkman, P.A. 2008. Once in a hundred generations. In: Halbert, M. and Skinner, K. (eds.). *Strategies for Sustaining Digital Libraries*. Emory University, Atlanta. Pp. 11–21. All rights reserved from EvREsearch

¹² Paul Berkman, 'Defining Digital Library Sustainability' (Paper presented at the Sustaining Digital Libraries Symposium, Atlanta, 6 October 2006)
<<http://www.metascholar.org/events/2006/sdl/viewpaper.php?id=6>>.

The progression from different technologies has the advantage of increasing transportability. The volume in which the information can be compacted and the density of information that can be transported is increased and the ability to integrate different types of information improves along the trajectory. These are all positive benefits; however there is an important retrograde projection: fragility.

People can still read a Gutenberg bible printed six centuries ago, but it can be challenging to read magnetic media a decade or two old. Fragility increases along the trajectory and this is a significant challenge for preservation. It predicates a fundamental paradigm change in preservation strategies. An example of the loss of important information is the first electronic mail message. This was sent in 1964 from either MIT, the Carnegie Institute or Cambridge University, however the message does not survive and there is no record to determine which group sent the first email.¹³ A less fortuitous example of loss is NASA losing more than 13 000 original tapes of the Apollo moon missions.¹⁴

A survey completed in 2006 by the United States National Library of Medicine found that of the 6 054 articles in 214 journal issues published in 2006, in the biomedical arena, 10% of the articles have linked digital information or supplementary digital information. What occurs to these links over time? Carmine Sellitto completed a study in 2004 (see Figure Seven) which showed that after just one year 10% of the links are broken.¹⁵ The half-life of links in the study was approximately four and a half years.

¹³ Report of the Task Force on Archiving of Digital Information (commissioned by The Commission on Preservation and Access and The Research Libraries Group, May 1996), 3 <http://www.digitalpreservation.gov/pdf/waters_garrett_final-report.pdf>.

¹⁴ Seth Borenstein, 'NASA Plans New Search for Missing Moon Tapes' *Houston Chronicle* (Houston) 15 August 2006 <<http://www.chron.com/dispatch/story.mpl/front/4116978.html>>.

¹⁵ Carmine Sellitto, 'A Study of Missing Web-cites in Scholarly Articles: Towards and Evaluation Framework' (2004) 30 *Journal of Information Science* 484 <<http://jis.sagepub.com/cgi/reprint/30/6/484>>.

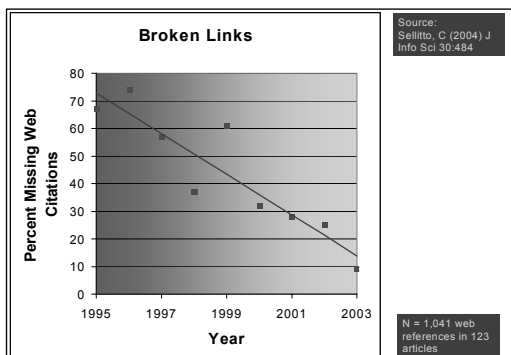
Figure Seven

Figure Eight¹⁶ is a compendium of similar studies. In 2002 legal citations were analysed and at the time were found to have a half life of less than one and a half years. The loss of information that has been published in the formal publication realm is significant and systematic and ranges in the analysis from one and a half years to four and a half years.

Figure Eight

Diagram omitted from digital version

Please see hard copy for diagram. For hard copy, please contact Sydney University Press (www.sup.usyd.edu.au)

Source: Koehler W. (2004) *Information Research*, 9 (2), 174

¹⁶ W Koehler, 'A Longitudinal Study of Web Pages Continued: A Consideration of Document Persistence' (2004) 9 (2) *Information Research* 174 <<http://informationr.net/ir/9-2/paper174.html>>.

There is an exception to the half life of citations as demonstrated in the studies. Digital objects that have been preserved in a formal digital repository, such as the Stanford Linear Accelerator Repository, the Harvard Digital Library and PubMed Central, have a half life of nearly 25 years. It could be argued that given the nature of these objects, this might be closer to the proper half life of digital objects. While some information should not be kept indefinitely, other information should be kept for a longer period of time. This illustrates the role of formal digital preservation organisations and their importance in the five dimensional world.

DATA PRESERVATION AND ACCESS: A SHARED VISION

A task force report issued in 1996 raises the challenge to ‘commit ourselves [as a society] technically, legally, economically, and organizationally to the full dimension of the task of preservation and access.’¹⁷ This is a fundamental challenge and progress has been made in globally recognising the nature of the challenge.

Organisation for Economic Co-operation and Development

The Organisation for Economic Co-operation and Development (OECD) believes that the issue of preservation of and access to research data is ‘a matter of sound stewardship of public resources.’¹⁸

Digital Repository Infrastructure Vision for European Research

The Digital Repository Infrastructure Vision for European Research (DRIVER) project is of the opinion that ‘any form of scientific-content resource . . . should be freely accessible through simple Internet-based infrastructures.’¹⁹

¹⁷ Commission on Preservation and Access and the Research Libraries Group, *Report on the Task Force on Archiving of Digital Information* (1996).

¹⁸ Organization for Economic Co-operation and Development, *Promoting Access to Public Research Data for Scientific, Economic and Social Development*.

¹⁹ Digital Repository Infrastructure Vision for European Research (DRIVER) <www.driver-repository.eu>.

Canada

The National Consultation to Scientific Research Data (NCASRD) acknowledges the importance of a robust infrastructure framework for digital preservation and access and proposes the ‘establishment of a dedicated national infrastructure . . . to assume overall leadership in the development and execution of a strategic plan [for digital data].’²⁰

New Zealand

Creating Digital New Zealand: The Draft New Zealand Digital Content Strategy emphasises the importance of preserving the digital products of the current culture for future generations and ‘providing the mechanisms to make it quick and easy. . . to find, share, access, use and re-purpose content.’²¹

Australia

The National Library of Australia’s Preserving Access to Digital Information (PADI) initiative aims to ‘ensure that digital information is managed with appropriate consideration for preservation and future access.’²²

ACHIEVING THE VISION

The National Science Foundation has a vision in which ‘science and engineering digital data are routinely deposited in well-documented form, regularly and easily consulted and analyzed . . . and openly accessible’ whilst being reliably preserved.²³

²⁰ The National Research Council Canada, *Final Report of the National Consultation on Access to Scientific Research Data* (2005) 3 <http://ncasrd-cnadrs.scitech.gc.ca/NCASRDReport_e.pdf>.

²¹ National Library, *Creating Digital New Zealand: The Draft New Zealand Digital Content Strategy Discussion Document* (2006) 7 <<http://www.digitalstrategy.govt.nz/upload/Main%20Sections/Content/NZ%20Digital%20Content%20Strategy%20Discussion%20Document.pdf>>.

²² Leanne Brandis and Jan Lyall, ‘PADI: Preserving Access to Australian Information and Cultural Heritage in Digital Form’ (Paper presented at the VALA Conference, Melbourne 28–30 January 1998) <<http://www.nla.gov.au/nla/staffpaper/lyall3.html>>.

²³ National Science Foundation, *Cyberinfrastructure Vision for 21st Century Discovery*.

There are three parts to achieving this vision. The first is that science and engineering data should be ‘routinely deposited in well-documented form’.²⁴ This is not a technology challenge, because the technology exists. Instead this is a cultural change for incentives and motivations to deposit and provide documentation for data.

Secondly that data should be ‘regularly and easily consulted and analyzed’ by specialists and non-specialists.²⁵ There are deep research and technology challenges to providing information accessibility for those who are not highly specialised in the field of that particular collection.

Thirdly, data should be ‘openly accessible while suitably protected, and reliably preserved’.²⁶

In order to achieve this vision, it is necessary to have an infrastructure framework: a framework of repositories, libraries and reliable preservation organisations to provide for this function. In order to meet this vision the National Science Foundation has set itself two goals. Firstly to catalyse the development of a system of science and engineering data collections that is open, extensible and evolvable. While the National Science Foundation cannot meet the digital preservation needs of society as a whole, it can play an important role in demonstrating this ability and establishing more appropriate methodologies and capabilities.

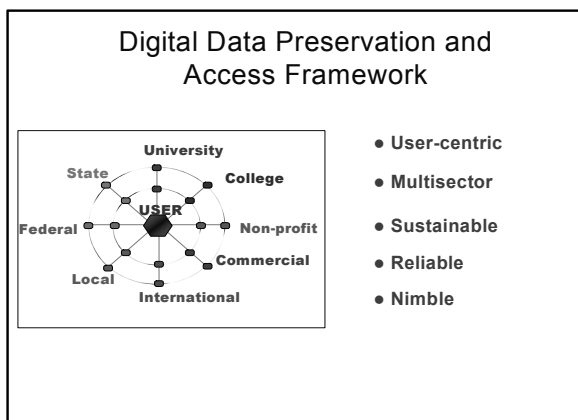
Secondly the National Science Foundation will need to develop new tools and servers to enable this infrastructure framework.

Figure Nine is a schematic of what the National Science Foundation envisions. At the centre of the schematic are the users using the infrastructure. The nodes represent individual repositories or digital libraries and the edges between the nodes represent the links between them.

²⁴ National Science Foundation, *Cyberinfrastructure Vision for 21st Century Discovery*.

²⁵ National Science Foundation, *Cyberinfrastructure Vision for 21st Century Discovery*.

²⁶ National Science Foundation, *Cyberinfrastructure Vision for 21st Century Discovery*.

Figure Nine

This schematic has several important features. Firstly the schematic is centred around the user and not around the infrastructure itself. The schematic occurs in sectors including federal, state, university, not-for-profit, commercial and international. The schematic should also be sustainable.

The schematic could be a schematic for the existing system of libraries preserving print information. There are international, national, local and university libraries which all have a different, but related, set of roles in preserving print information. These libraries also have a variety of business models through which they draw their funds from a variety of different sources in society.

The net result is a system which is robust and resistant to change in any one sector or catastrophic loss. This is the type of multi-sector sustainable framework that the National Science Foundation envisions. The framework should be reliable and the metrics for reliability and the technologies for reliability are important and are still being developed. The framework will also have to be nimble, because it operates in a swift current of constant technology change.

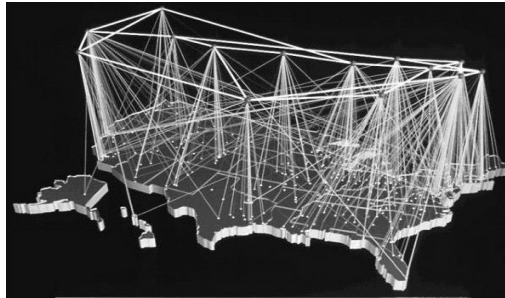
In summary the National Science Foundation's strategic plan includes promoting a change in culture, developing the preservation framework

and supporting the new generation of tools, services and capabilities that this framework will require.

Figure Ten is a graphic from the National Science Foundation summarising traffic on the NSFNet in September 1991. The NSFNet was arguably one of the best infrastructure investments the National Science Foundation made. NSFNet was the consolidation of two precursors, the ARPANet (the Defence Agency Network) and CSNet (the Computer Science Net) which were consolidated in order to provide access to the newly launched supercomputer centres in the United States.

Figure Ten

NSFNet Traffic– September 1991



Source: Visualization prepared by NCSA using data provided by Merit.

NSFNet was intended as an academic network. When it was launched in 1986 the National Science Foundation made the then outlandish claim that in five years time the NSFNet would connect up to 200 academic institutions with 10 000 users, which at the time seemed to be an immense goal. But by the end of 1992, when the NSFNet T1 net was decommissioned in favour of the T3 network there were one thousand institutions connected to the NSFNet with 10 million users.

Opening NSFNet to everybody resulted in the growth of the Internet. The Internet created connectivity, the ability to connect one machine to another without necessarily having to know in advance where that machine was located. It was that simple power of connectivity that

drove the emergence of the infrastructure and the opening of the fifth dimension.

The next driving force of this type is information integration, which requires reliable preservation. To bring about information integration, the National Science Foundation may have to start out relatively small with an initial datanet that is fairly simple and link together a small number of repositories. This will demonstrate the power of access to a wide variety of information and the ability to integrate that information. Then, if the force of integration is analogous to the force of connection, it is possible that after a short period of time, the datanet will grow exponentially.

There are two entities, or types of organisations, that are critical in this next stage. The first are the universities. Andre Oosterlinck states that the traditional function of the university is to create knowledge through research, disseminate knowledge through teaching and public outreach and preserve knowledge through the library systems of the university.

Ever since their inception, universities have been occupied with the fundamental elements of what we now call 'knowledge management', i.e. the creation, collection, preservation and dissemination of knowledge.²⁷

This responsibility is reflected in the mission statement of the University of California.

The distinctive mission of the University is to serve society as a center of higher learning, providing long-term societal benefits through transmitting advanced knowledge, discovering new knowledge and functioning as an active working repository of organized knowledge.²⁸

The universities are in a unique position. The mission of universities is consistent with the affirmation of the shared vision mentioned above. Some universities and libraries are amongst the oldest organisations in the world. The universities are organisations that have substantial information technology capabilities and faculties that generate digital

²⁷ Andre Oosterlinck, *Knowledge Management in Post-Secondary Education: Universities.*(2002)

²⁸ University of California, *Mission Statement.*

data and computer science breakthroughs, which are cyberinfrastructure advances that are critical to the evolution of the datanet concept.

Similarly the academic libraries have an important role to play. The Association of Research Libraries²⁹ is a group of 123 North American academic libraries whose mission is the preservation of digital assets.

It is to the research library community that others will look for the preservation of . . . digital assets, as they have looked to us in the past for reliable, long-term access to the 'traditional' resources and products of research and scholarship.³⁰

The University of Queensland library envisions a similar role in providing a link between people and information:

The University of Queensland Library's mission is to link people with information, enabling the University of Queensland to achieve excellence in teaching, learning, research, and community service.³¹

I-Centre

However the current structure of the university and the university library is not optimal for the access to and preservation of digital information, and a new type of organisation is necessary. For the purposes of discussion this organisation will be called an 'I-Centre'.

It is necessary for the I-Centre to be risk-averse. It must have a timeline for reliable preservation of digital content that stretches into centuries while anticipating how people in the future will use the information that has been preserved. The expertise necessary for developing the I-Centre and its risk-averse capabilities lies in the library and archival sciences.

At the same time the hardware, software and people at the I-Center will be changing. Therefore the organisation has to be risk-capable, it has to be able to operate within a swift current of constant technology change and a steady exponential increase in the expectations of user who will

²⁹ Association of Research Libraries <<http://www.arl.org/>>.

³⁰ Association of Research Libraries, *ARL Strategic Plan 2005–2009* <<http://www.arl.org/arl/governance/stratplan.shtml>>.

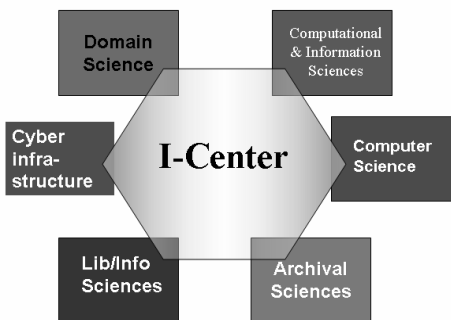
³¹ Keith Webster, University Librarian and Director of Learning Services.

want more from the cyberinfrastructure than what it is attempting to deliver.

For this reason, the organisation must have capabilities in computer science and computational science to anticipate the next generation of technologies, identify risks associated with those new technologies and plan reliable migration to the new technologies. This will be a constant occurrence through the life of the I-Centre.

Finally the user must be able to understand and access the information. This will require domain expertise necessary for understanding the deep contextual information associated with the information being preserved. An understanding of how the information will be used in the community will also be necessary, and this will require significant expertise in the respected domains.

Figure Eleven



The I-Centre is an organisation that for the most part does not exist. A change may be required in the nature of digital preservation organisations, which will require new partnerships that are currently not present.

AN EFFECTIVE LEGAL FRAMEWORK

The vision of this conference, for an effective legal framework raises some fundamental issues.

The end result of the framework is information integration, everything else is the means towards achieving this. The goal of this cyber-

infrastructure framework is the ability to find, understand, access, use and re-use information. Any legal framework that inhibits or prevents information integration will inhibit the progress of those who operate under that framework.

The foundation of the framework has to be reliable digital preservation and access. If information is constantly lost, not accessible, or moving and changing in significant ways, the ability to effectively use the information over time is significantly decreased.

There are many types of data, for instance data that are public goods and data that are commercial commodities. An effective legal framework must recognise and support the various types of data. The legal framework should not focus on one single category of data, or a finite set of categories into which the data types can fit over time. Rather the framework should recognise the many different types of data being produced and the many different uses and needs for that data.

A world of five dimensions is inherently international, not national in character. In science the closest alignments between individuals are within disciplines, not within geographical regions. People operating in a world of five dimensions will be operating in an international framework. While this is a given for the five dimensional world, it must be an essential part of an effective legal framework.

It must be recognised that the fifth dimension does not arise automatically; it is built by individuals and organisations. The framework should enable individuals and institutions to pursue their innovative approaches to the infrastructure of the future.

Finally, there is a constant change in technologies, users' needs and expectations and opportunities. The legal framework must be built on the assumption that a static framework is dangerous and will almost certainly break immediately. The ability to accommodate a continuing change in the technologies is critical, and failing to do this will put the system at risk. The technology in this area will always improve and the legal framework should anticipate continuing change in this landscape.

CONCLUSION

The Office of Cyberinfrastructure is currently working on the technology challenges and opportunities that exist in creating this fifth dimension. In doing so it has been recognised that it is important to

have an adequate and robust legal framework to enable the technology innovations that are necessary. The legal work being done in this area is just as critical to the future of the five dimensional world as the technology that is being created.

INNOVATION AND OPEN ACCESS TO PUBLIC SECTOR INFORMATION

Dr Terry Cutler¹

Most speakers at this summit have been looking at open access from the supply side, presenting the points of view of custodians of government information. What might we lob over the fence to whoever is on the other side? So far we have not paid much attention to this demand side - the potential beneficiaries of changed information policies. So I see it as my task to address what I believe is the core rationale for this policy initiative, which is the promotion of innovation and creativity. My perspective on the topic brings together my deep interest in the whole matter of innovation, and my long involvement with the digital content industries.

Why do we need to act on this possible policy initiative? I will try to put the question in the context of some conceptual frameworks and models of innovation, and of business models for information and content production. My premise is that data and information – content – is the currency of creativity and innovation. Information is what energises our national innovation system. Governments produce and hold a wealth of information and data.

Both creativity and innovation have become somewhat fuzzy terms. This leads me to begin with two texts for today, one secular and one sacred (in the interests of balance and even handedness). My first text

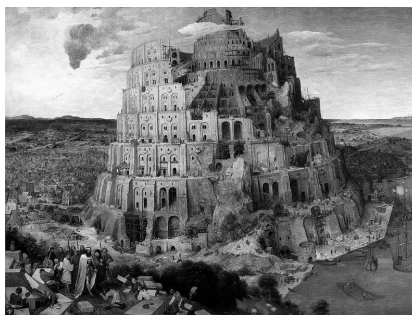
¹ Principal, Cutler & Co. This chapter is an extended version of speaking notes from the Australian National Summit on Open Access to Public Sector Information convened by the Law Faculty of Queensland University of Technology and supported by the Queensland Spatial Information Council. The Summit was held in Brisbane, Australia, on 13 July 2007, the day following the *Legal Framework for e-Research Conference*. I have taken the opportunity to elaborate upon my presentation in the interests of clarity. This paper draws on other work in progress, and my 2006 submission to the Productivity Commission's Inquiry into Public Support for Research and Innovation.

comes from the venerable Henry Fowler's *Modern English Usage*, where he writes:

creative is a term of praise much affected by the critics. It is presumably intended to mean something original, or something like that, but is preferred because it is more vague and less usual (cf. *Seminal*). It has been aptly called a 'luscious, round, meaningless word', and said to be 'so much in honour that it is the clinching term of approval from the schoolroom to the advertiser's studio'.

In other words, Fowler finds our use of the term 'creative' just a little bit vacuous. It's probably fortunate he died before the word 'innovative' became the new 'clinching term of approval'. Now many of those working on this open information initiative are lawyers, and what I like and respect about lawyers is their precision about words and terminology. The construction of language is at the core of their craft, and we can usefully apply this rigour to the reconstruction of meaning around innovation and creativity.

I take my second text from Genesis, and the account of the destruction of the tower of Babel. The Tower of Babel provides us with a splendid metaphor for the creation of a perfect market in information. For those who may have forgotten how the story goes, let me remind you of the text and try to draw out the lessons for today.



Source: Wikipedia,
<http://en.wikipedia.org/wiki/Image:Brueghel-tower-of-babel.jpg>

Artist: Pieter Bruegel c. 1525/30

Genesis Chapter 11 begins with a vision of an information paradise (and a vision of ‘whole of government’ coherence) – ‘*one language and one speech*’.

1 And the whole earth was of one language, and of one speech.

2 And it came to pass, as they journeyed from the east, that they found a plain in the land of Shinar; and they dwelt there.

3 And they said one to another, Go to, let us make brick, and burn them thoroughly. And they had brick for stone, and slime had they for mortar.

4 And they said, Go to, let us build us a city and a tower, whose top may reach unto heaven; and let us make us a name, lest we be scattered abroad upon the face of the whole earth.

5 And the Lord came down to see the city and the tower, which the children builded.

6 And the Lord said, Behold, the people is one, and they have all one language; and this they begin to do: and now nothing will be restrained from them, which they have imagined to do.

7 Go to, let us go down, and there confound their language, that they may not understand one another's speech.

8 So the Lord scattered them abroad from thence upon the face of all the earth: and they left off to build the city.

9 Therefore is the name of it called Babel (confusion); because the Lord did there confound the language of all the earth: and from thence did the Lord scatter them abroad upon the face of all the earth.

- *Genesis 11:1–9*

Verse 6 reminds us of the power of a common infrastructure and shared knowledge. But suddenly, in the following verse, what I will render as ***Adam Smith's curse*** descends on us. All that ‘*which they have imagined to do*’ is struck down through the specialisation of labour, the segmentation

of academic disciplines and discourse, and the bureaucratisation of governance. Fragmentation and confusion ensues.

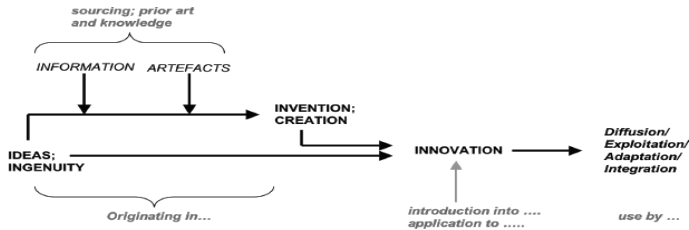
But, at the end, there is hope. An unintended consequence of the destruction of the Tower of Babel is the creation of diversity. And diversity is widely recognised as a pre-condition for creativity and innovation.

I will labour the point about the importance of precise language and clarity about concepts like innovation and creativity because, otherwise, these terms do not serve as reliable guides to action. We also need to remind ourselves regularly of just why being innovative is so important.

Innovation is critical to the competitiveness and sustainability of our economy and society. Yet, for all the fuzzy talk about it, and for all the platitudinous reports and business school prescripts, it is rarely the subject of rigorous examination and critical thinking. It is difficult to find a coherent, comprehensive account of innovation. You will find it difficult to unearth the term in standard economic textbooks. The reason for this is because neo-classical economics works predominately with closed models of the market: equilibrium models. Innovation, however, is all about change and economic development: disequilibrium and the breakthrough thinking from which we learn and build our stock of knowledge and, hopefully, of wisdom.

To set out an account of innovation I need to begin with a taxonomy of the terms involved, and the related concepts. With such building blocks we can begin to explore the dynamics of innovation as a change and learning process.

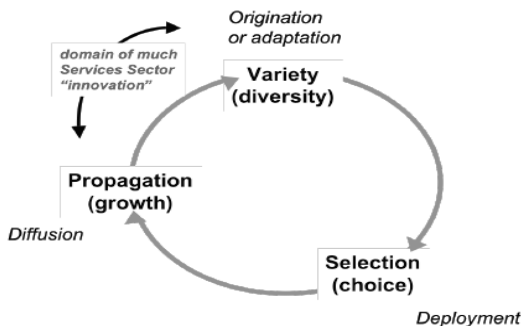
INNOVATION: A TAXONOMY OF TERMS AND RELATED CONCEPTS



Source: Dr Terry Cutler, Submission to the Productivity Commission Inquiry into Public Support for Science and Innovation, July 2006, p7, available at www.pc.gov.au/__data/assets/pdf_file/0006/37662/sub043.pdf

Simply linking these terms and then sequencing them according to the underlying grammar – analogous to a DNA sequence – we can begin to derive a theory of innovation.

INNOVATION AS AN OPEN SYSTEM



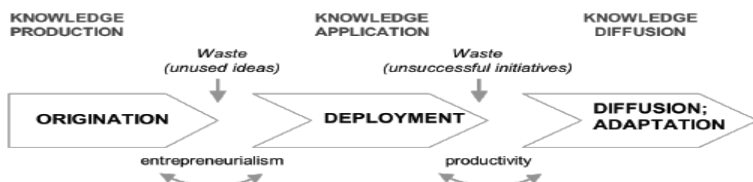
Source: Dr Terry Cutler, Submission to the Productivity Commission Inquiry into Public Support for Science and Innovation, July 2006, p11, available at www.pc.gov.au/__data/assets/pdf_file/0006/37662/sub043.pdf

This is an open model -as distinct from the closed models of neo-classical economics -which is comparable to and, indeed, refers to the open models we find in the life and physical sciences.²

The energy in this open system of innovation is creativity: the ideas and insights which produce the options for doing something differently. The accumulation of such thinking is a pool of options for future development. Without new ideas, this pool is not refreshed and becomes stagnant.³ The value of ideas and inventions only comes into play when they are applied to problems or opportunities in markets or the community. The value is only fully realised when the innovation is taken up and used widely. In the process of adopting an innovation, moreover, adaptations and improvements will occur. This is because adoption will normally require adaptation to the context of the use. Thus the open-ended cycle of change and renewal will continue.

To elaborate this model into a more fully rounded theory of innovation we need resort to a mercator-like projection of the schematic.

THE MERCATOR PROJECTION OF INNOVATION THEORY



Source: Cutler 2006

² As an elaboration – beyond the confines of a short speech – this theoretical model resonates with Darwinian exposition and the language of thermodynamics and negative entropy. It is worth observing that Adam Smith himself would not be alarmed; Smith's whole opus shows an acute awareness of historical progression and tipping points. His successor, Alfred Marshall, notably regretted never returning to the bigger picture of the dynamics of political economy after his excursion into the domain of abstracted and closed economic models. Schumpeter famously took up the challenge, but never quite got there.

³ This insight is about how we actually can go backwards, as the history of many cultures demonstrates. This is the entropy of knowledge.

There are five points I want to draw out from this schematic.

The first is that creativity and fresh thinking is invoked within each of the sub-systems or elements of the innovation process, and at the points of intersection between them.

Secondly, we need to look carefully at the entrepreneurial process of matching a capability with a need or opportunity. This is a purposeful process of selection, not a linear progression of ideas simply walking out of the laboratory or study into the marketplace. It is more productive to seek solutions to a need or opportunity than to hawk solutions in the search for a problem. This observation is, of course, at odds with contemporary cargo cults about the commercialisation of research.

Thirdly, productivity arises from the successful deployment of innovations, not from the innovation per se.

Fourth, information and data is the basic currency across this whole ecosystem.

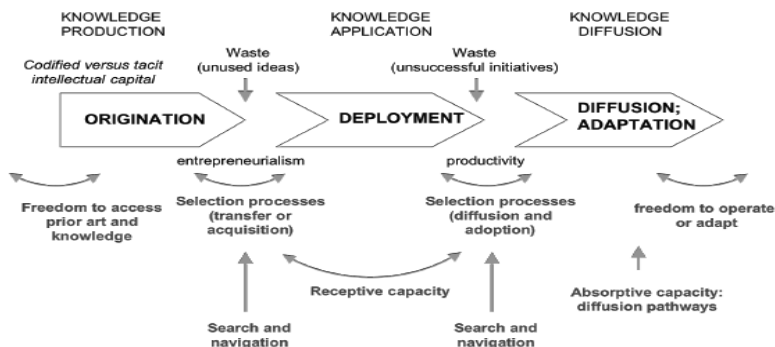
Fifth, there is waste in the system, whether unused ideas – including possibilities stored away for revisiting later or ideas whose time has not yet come – or failed ventures, including situations where a venture may fail for reasons other than the merit of the innovation.

These last two points are highly relevant to considerations around access to public sector information. The originator, owner or custodian of information or data may not be best placed to understand the possible uses or potential future uses of the information or data they hold. Waste and the destruction of value may occur because government sets rules of access to information which fail to recognise the requirements of unforeseen users and uses. Furthermore, the rules of engagement between government and the initial agent – the immediate user or use – may unintentionally constrain the beneficial use by third parties or eventual end-users in the process of the diffusion of knowledge or innovation.⁴

While information is the currency of innovation, informational and content sources play different roles within different parts of the innovation system.

⁴ Examples are pricing models or the processes of access, including technological requirements. Restrictions on 'primary data' or source code may inhibit useability and re-use.

ACCESS ISSUES AROUND INFORMATION AND KNOWLEDGE



Source: Cutler 2006

Knowledge builds on knowledge. This has some important implications for innovation, and for considerations of access to public sector information.

First, even when an entrepreneur sees an opportunity, they need certain skills and domain knowledge to be able to understand the potential of new ideas and knowledge and to act on the opportunity. The existence of such skills will affect the capacity to form effective collaborations, whether as a firm or a project. The innovation process will falter in the absence of effective partners or collaborators. We talk about this as the receptive capacity of an industry or body politic. There may often be a public policy interest in improving this receptive capacity. Without it, innovation will be constrained.

Secondly, the wider diffusion and take-up of an innovation depends on the absorptive capacity of the community. For example, the take-up and sustainability of certain information technologies requires particular skill levels within the user population. Data sets are meaningless without the requisite analytical skills. Thus the education and skill levels of the general population become important considerations for everyone.

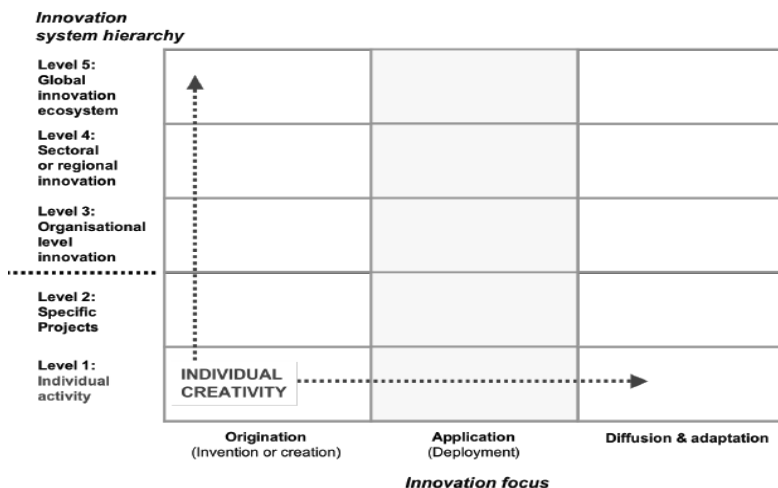
Thirdly, certain freedoms are essential to creativity and innovation.⁵ The first is the freedom to access and use prior art and knowledge in the exploration and development of new knowledge and insights. It is obvious that open access underpins this freedom. Equally important, however, is the freedom to operate and adapt in the process of deployment and diffusion. The extent of this freedom will depend on what rules and conditions are imposed by the owners of an innovation. The terms of access to information and data will dictate the extent of further experimentation and development. This becomes particularly important when an innovation can usefully be packaged or integrated with other products or services. Systems integration is an increasingly significant platform for innovation, especially in the services sector.

My final point about innovation is that it is a complex system. Innovation functions at multiple, interdependent levels. At the heart of the matter is the individual person: call them artist, scientist, technician, knowledge worker or whatever. Individual people fuel the whole innovation system. We also talk a lot about collaborations between people, but for all the rhetoric we know that in practice it is hard.

The following matrix identifies five levels within an innovation system, each with discrete issues but all are highly interdependent. For each level there are discrete and distinctive institutional and human capital issues to be taken into account with each element of the innovation process (of origination, deployment, diffusion and adaptation).

⁵ Both Karl Popper and Amartya Sen should be essential reading for any naysayer.

THE MATRIX OF INTERACTIONS WITHIN THE INNOVATION SYSTEM



Source: Dr Terry Cutler, Submission to the Productivity Commission Inquiry into Public Support for Science and Innovation, July 2006, p18, available at www.pc.gov.au/__data/assets/pdf_file/0006/37662/sub043.pdf

Issues around access to information and content, and the role of public sector information, will vary across this matrix, both horizontally and vertically. It is arguable that simple and flexible digital content architectures will maximise the utility of public sector information and data sets across the variety of user environments implied by this matrix. The principle should be to empower the greatest possible range of uses, known and unforeseen.

I have argued that innovation is an open system. This resonates with industrial firms who increasingly are paying attention to the flow of knowledge and intellectual capital across organisational boundaries.

For most of the twentieth century firms pursued a model of in-house, proprietary research and development to sustain their innovation. With globalisation and the deconstruction of supply chains this model has become unsustainable. The dominant model of innovation has changed to an open model drawing on multiple internal and external sources of

ideas and channels to market. This open innovation model emphasises *knowledge flows* rather than knowledge creation as a driver of innovation.

THE KNOWLEDGE LANDSCAPE IN THE OPEN INNOVATION PARADIGM

Diagram omitted from digital version

Please see hard copy for diagram. For hard copy, please contact Sydney University Press (www.sup.usyd.edu.au)

Source: From *Open Innovation: The New Imperative for Creating And Profiting from Technology* by Henry William Chesbrough, pp47. Copyright © 2003 by the Harvard Business School Publishing Corporation; All rights reserved. Reprinted by permission of Harvard Business School Press

Open innovation models recognise that one person's trash is another person's treasure. There is an inbuilt asymmetry between the owners or custodians of information, and potential users in terms of the uses of information and the value of those uses.

At a conference on e-Research which preceded this Summit, Dr Chris Greer from the National Science Foundation in the US spoke of cyberinfrastructure as a new fifth dimension and shared space. In thinking about such information and collaboration infrastructures, it struck me that many of the access issues we are debating around digital information have already been addressed in other domains, especially around open access to physical infrastructure. There are clearly lessons to be learned from the principles established for access to and the interconnection of deregulated telecommunications networks, and other forms of networked infrastructure.⁶

⁶ I was personally involved in the early debates on these issues during the liberalisation of telecommunications markets in Australia and Asia in the early 1990s. Much of the clarity and sharpness of the principles then established has been eroded over time.

Access regulation for telecommunications networks is based on two major premises:

1. The utility and benefits of networks are promoted by ‘any to any’ connectivity (interoperability); and
2. Dominant players should not be able to create ‘bottlenecks’ to access.

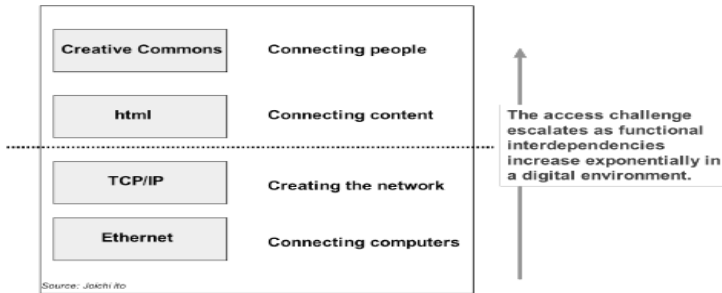
A number of access principles⁷ follow from these premises and include:

- Arrangements should promote efficiency;
- There should be reciprocity in rights and obligations;
- The economics of arrangements should be clear and unbundled, promoting:
 - The desired level of investment in infrastructure (without wasteful duplication)
 - The lowest possible transaction costs;
- Obstacles to users accessing services should be minimised; and
- Redundancy should be supported.

Network ‘interconnection and access’ principles are clearly applicable to information infrastructures and content networks. Content is the new access bottleneck. The access challenge escalates as functional interdependencies increase massively in a digital environment. As a principle, networked information flows should aim to support ‘any to any’ connectivity. This seems especially apposite in the case of public sector information.

⁷ Australian Telecommunication Authority (AUSTEL), *Study of Arrangements and Charges for Interconnection and Equal Access*, (1991).

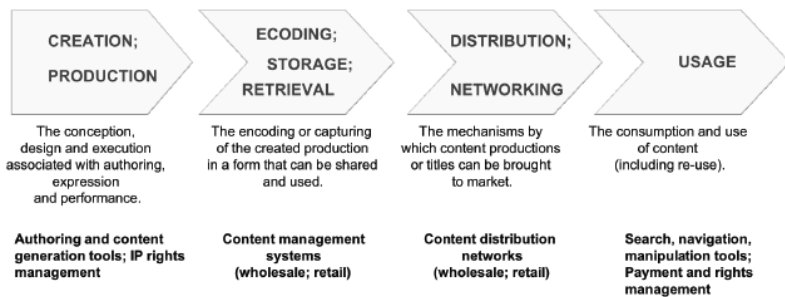
CONTENT AND INFORMATION ARE THE NEW AREAS OF ACCESS BOTTLENECKS



Source: Joi Ito (2006).

Developing policies on access to information requires attention to the whole business system of content and information production. In a digital environment, the business system of content revolves around bit creation, bit storage, bit distribution, and bit use and re-use. A model I developed around this in 1994 still seems to stand up:

THE BUSINESS OF DIGITAL CONTENT

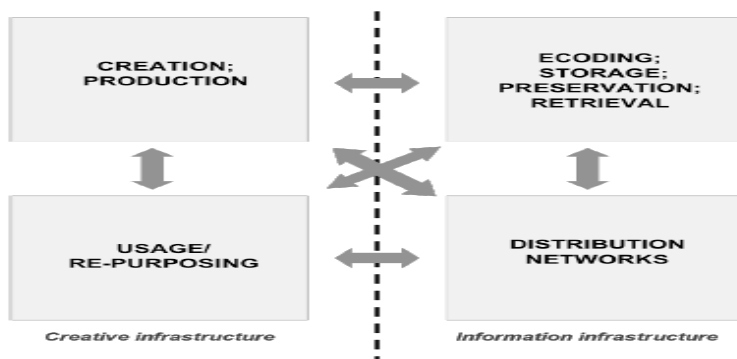


Source: Cutler & Company, *Commerce in Content* (1994),
<http://www.nla.gov.au/misc/cutler/cutlercp.html>

In management jargon, digital content (information) production is more of a 'value net' than a serial value chain. This is because of the functional interdependencies within a digital ecosystem. The 'freedom to

operate' and create within the producer or user environment – include the re-purposing of information – will be facilitated or constrained by the functionality of the supporting information infrastructure and its architecture. Policies for open access need to minimise the obstacles which may arise from these functional interdependencies.

THE DIGITAL CONTENT ECOSYSTEM



Source: Cutler & Company, *Commerce in Content* (1994),
<http://www.nla.gov.au/misc/cutler/cutlercp.html>

Why is open access to public sector information important for innovation? I have argued that it is important because knowledge and information flows underpin creativity and innovation. It is especially important in a small country economy like Australia because of the relative scope and scale of public sector information. The public sector is a major – even the dominant – producer and custodian of information. Furthermore, only government and the public sector have the critical mass to create inclusive public platforms and scalable repositories.

Ironically, open access policies could also help resolve the chronic problems with 'silo' barriers to information sharing *within* government – promoting greater 'whole of Government' effectiveness.

CONCLUSION

Information infrastructure and information architectures are crucial in an information society. Government information policies should promote:

- ‘freedoms to operate’ – ‘unfreedoms’ are the enemy of development and innovation; and
- open, end-to-end access as a fundamental premise of infrastructure.

The wise administration of public sector information can create significant economic benefits through strengthening the national innovation system. By its own practice, governments can help shape the rules and conduct of wider information markets. As with most things, however, the devil is in the detail. The utility of public information to users will be determined by the terms of access, including the efficacy of arrangements for such things as:

- information exclusions – open access should be the default setting;
- searchability and discovery;
- transparency of language and code;
- transaction costs; and
- the preservation of information and its long-run accumulation.

Good outcomes will require us to approach the principles of access from the perspective of prospective users, and with a keen regard to the potential obstacles and bottlenecks to the effective use of public sector information.

CYBER INFRASTRUCTURE FOR THE HUMANITIES AND SOCIAL SCIENCES

John Unsworth¹

In January 2003, a blue ribbon panel appointed by the National Science Foundation and led by Dan Atkins, of the University of Michigan, completed a report called ‘Revolutionising Science and Engineering through Cyberinfrastructure’.² This report is a kind of provocation for the American Council of Learned Societies (ACLS) Commission on Cyberinfrastructure for the Humanities and Social Sciences, and there is a lot of other activity of this sort, going on right now—for example:

- Digital Archiving and the National Archives and Records Administration;³
- NSF ‘Post Digital Library Futures’ report;⁴
- NRC ‘Beyond Productivity’ report (2003);⁵
- The United Nations World Summit on the Information Society.⁶

¹ Chair, Commission on Cyberinfrastructure for the Humanities and Social Sciences, American Council of Learned Societies; Dean and Professor, Graduate School of Library and Information Services, University of Illinois at Urbana-Champaign.

This chapter was first presented as a paper at the annual meeting of the Research Libraries Group (Washington, DC, 26 April 2004).

² Daniel Atkins et al, *Revolutionizing Science and Engineering through Cyber-infrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure* (January 2003) <http://www.communitytechnology.org/nsf_ci_report/>.

³ Digital Archiving and the National Archives and Records Administration (2004) <http://www7.nationalacademies.org/cstb/project_nara.html>.

⁴ National Science Foundation, *Knowledge Lost in Information: Report of the NSF Workshop on Research Directions for Digital Libraries* (2003) <<http://books.nap.edu/books/0309088682/html/index.html>>.

⁵ National Research Council of the National Academies, *Beyond Productivity: Information Technology, Innovation and Creativity* (2003) <<http://books.nap.edu/books/0309088682/html/index.html>>.

In a press release that followed the publication of the Atkins report, Peter Freeman, the assistant director of Computer and Information Sciences and Engineering (CISE), the NSF directorate that commissioned the report, said: ‘The path forward that this report envisions. . . truly has the potential to revolutionise all fields of research and education.’ Certainly, the report has had a significant impact on the rhetoric, and perhaps also on the priorities, not only of CISE, but also of other parts of NSF, and on other funding agencies concerned with information technology as it supports research.

So, what *is* cyberinfrastructure? Here is how the Atkins report addresses that question:

The term infrastructure has been used since the 1920s to refer collectively to the roads, power grids, telephone systems, bridges, rail lines, and similar public works that are required for an industrial economy to function. Although good infrastructure is often taken for granted and noticed only when it stops functioning, it is among the most complex and expensive thing that society creates. The newer term cyberinfrastructure refers to infrastructure based upon distributed computer, information and communication technology. If infrastructure is required for an industrial economy, then we could say that cyberinfrastructure is required for a knowledge economy.⁷

Cyberinfrastructure then is the infrastructure for a knowledge economy. And why should we care about it? Well, we all live, and will continue to live, in that knowledge economy, so we all have at least the same interest we would have in good roads and bridges, good telephone systems and power grids. And why should the humanities and social sciences care about it? Because we can make it a better infrastructure, if our perspectives, our training, and our expertise are included in its design and deployment. After all, science—whose goal is predictive certainty—only has half the picture. Uncertainty (or ambiguity, if you prefer) is the

⁶ United Nations, *World Summit on the Information Society* (2003, 2005)
<<http://www.itu.int/wsis/>>.

⁷ Daniel Atkins et al, *Revolutionizing Science and Engineering through Cyber-infrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure* (January 2003) 5
<http://www.communitytechnology.org/nsf_ci_report/>.

other half, and the humanities and social sciences celebrate that, explore it, tolerate it, and *understand* it better than the sciences do. Or, at another level, if science and engineering are about what we can do, the humanities and social sciences are about what we should do. If we do not know what we can do, we do not know what choices to consider, but if we do not know what we should do, we do not know which choices to make. Cyberinfrastructure is no different, in that respect, from atomic energy, biotechnology, or any other challenge: it is not only a scientific challenge, with scientific outcomes: it is also a social and human challenge, with outcomes that the humanities and social sciences are best equipped to understand.

The ‘overarching finding’ of the Atkins report

... is that a new age has dawned in scientific and engineering research, pushed by continuing progress in computing, information, and communication technology, and pulled by the expanding complexity, scope, and scale of today’s challenges. The capacity of this technology has crossed thresholds that now make possible a comprehensive ‘cyberinfrastructure’ on which to build new types of scientific and engineering knowledge environments and organisations and to pursue research in new ways and with increased efficacy. Such environments and organisations, enabled by cyberinfrastructure, are increasingly required to address national and global priorities, such as understanding global climate change, protecting our natural environment, applying genomics-proteomics to human health, maintaining national security, mastering the world of nanotechnology, and predicting and protecting against natural and human disasters, as well as to address some of our most fundamental intellectual questions such as the formation of the universe and the fundamental character of matter.⁸

I agree with all of this, and I am certain that Dan Atkins, and many other scientists, would agree that along with all this new knowledge, these new certainties, will come new uncertainty, and new quandaries, that science

⁸ Daniel Atkins et al, *Revolutionizing Science and Engineering through Cyber-infrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure* (January 2003) 31
<http://www.communitytechnology.org/nsf_ci_report/>.

itself, by itself, will not be able to resolve. But if the humanities and social sciences want to have some influence in the process now underway to design our information technology environment over the next decade, then we need to articulate our needs and our potential contributions—and even more than that, we need to articulate the importance of the humanities and the social sciences, for the amelioration of the human condition. That is something we have not done very well since progress displaced enlightenment as our culture's highest value, and it is something that we still do not do very well.

For example, in the 'Summary of 2003 Fiscal Year Budget Request', the NEH argues for the humanities as follows:

In the 1965 legislation that established the National Endowment for the Humanities, the Congress of the United States declared that 'democracy demands wisdom and vision in its citizens' and posited that 'promoting progress in the humanities' was the surest route to such wisdom ... The National Endowment for the Humanities helps Americans develop wisdom and vision through the study and contemplation of the record of human thought. The study of history, literature, languages, philosophy and other humanities subjects help us not only to better understand our own nation, but other cultures as well.⁹

All well and good, and I believe also true, but unfortunately we in the United States seem to believe that wisdom is a great deal less expensive than knowledge, and knowledge—especially knowledge with practical consequences—is what we are willing to spend money on.

But the humanities and the social sciences have access to knowledge that does have enormous practical consequence, and the future will be better, or worse, depending on whether that knowledge is part of our 'knowledge society'. In order to be a meaningful part of that society, I would argue, the humanities and social sciences will need computational methods and they will need access to the kind of vast datasets that make computational methods both necessary and useful. Computational methods already have a place in the social sciences, and they have a foothold now in the study of literature, history, art, and other humanities

⁹ National Science Foundation, *Summary of FY 2003 Budget Request to Congress*
<http://www.nsf.gov/bio/budget/bio_bdg03/bio_bgt03_toc.htm>.

disciplines. The humanities, in particular, has been without a galvanising methodology for a generation now, and it is being, and will be, revolutionised by information technology as profoundly as any of the sciences.

That revolution, though, has some preconditions: it requires a *motivating factor*, to move the disciplines toward new methods, and it requires the *means*—both intellectual and financial—to adopt, refine, and disseminate those methods for the rising generation of scholars.

The motivation, if it comes, will come in the form of very large datasets that can only be manipulated and interpreted with the aid of computers. We are getting there, already, with digital libraries, and if projects like DLF's distributed digital library come about, we will seem suddenly to have arrived. But even that scale is not quite what I have in mind. To arrive at terabytes or petabytes of humanities or social science data, we will have to effectively address two issues—intellectual property (the primary data-resource-constraint in the humanities) and privacy rights (the primary data-resource-constraint in the social sciences). In a sense, as John King pointed out to me, this is no different from the struggle, in computational science, with resource constraints on memory, bandwidth, or processing speed: unless we radically increase these resources, we radically limit the kinds of questions we can ask and answer. We generally think of intellectual property and privacy rights as legal issues, which of course they are, but short of sweeping legal remedies, which I do not expect and do not actually desire in either case, I think the solutions to these problems will be, to some significant extent, technical. That makes them, in effect, the primary 'cyberinfrastructure' research agenda for the humanities and social sciences. If we tackle that agenda successfully, the humanities will have access to the full recorded history of the 20th century—music, film, text and image, all in digital form; the social sciences will have access to the full record of societies, populations, individuals. There is tremendous danger of abuse, here—as there is with any other research that has profound practical consequences—but there is also a tremendous opportunity to learn, to understand, even to achieve some wisdom. Lest that emphasis on big datasets sound too much like 'rugged informationalism', let me emphasise that I agree with David Weinberger that 'what the world needs [is] people who know how to manage metadata, navigating the twisty darkness of the ambiguous world while

preserving the value of the unspoken.’ I would just add that if we only have access to metadata, and not to the data itself, we will not be able to really plumb ambiguity’s ‘twisty darkness’— but I would also agree that ‘preserving the value of the unspoken’ is one of the major challenges for the humanities, especially, as it grapples more deeply with computation and information science.

If you look at where the funding for cyberinfrastructure and the research programs that support it is coming from, you might conclude that not all communities, not all classes, and not all categories are going to be equally well served in its design and deployment. To begin with, and lest we lose sight of this, the largest investment will be from the commercial sector, and it will take the form of developing products, not doing basic research, much less doing education. A knowledge society implies an information economy, and we have already seen that the owners of large caches of information—television and film studios, the recording industries, publishers—are not eager to achieve technical solutions to the problem of restricted access.

In the area of basic research, most of which is now done in universities and colleges, the big dog is health research, which accounts for more than half of federal spending on research: the NIH’s budget in 2004 was around \$28B. By comparison, the NSF’s budget was about \$5.5B (an increase of \$171M over 2003), within which CISE represents about ten percent, or \$584M. By comparison, again, \$584M is about what the largest private foundations give away in a year, in this country—with the exception of the Gates Foundation, which gives away about twice that. Descending the scale of funding and influence, the 2004 budget for IMLS was about \$262M (roughly half of the budget for CISE); the budget for NEH was \$162M (less than the increase for CISE, over 2003); bringing up the rear, the budget for NEA was \$139M—just under half a percent of the budget for NIH. In fact, add the budgets of NEA, NEH, and IMLS together, and you will not quite equal the budget for CISE, which is one of the mid-range budgets in NSF.

‘Cyberinfrastructure’ is more than just hardware and software, more than bigger computer boxes and faster wires connecting them. The term describes new ‘research environments’ in which disciplinary experts, in interdisciplinary teams, supported by specialised computational support staff, have global, instantaneous access to enormous computing resources. And although the redaction of the Atkins report in NSF

presentations subsequent to its publication has tended, in my view, to emphasise only that last point—enormous computing resources—if you read the report itself, you may be struck, as I was, by its emphasis on human resources, on organisations, and on education and training, for example in passages like this one:

This [vision of science and engineering research] involves significant educational dimensions. The research community needs more broadly trained personnel with blended expertise in disciplinary science or engineering, mathematical and computational modelling, numerical methods, visualisation, and the sociotechnical understanding about working in new grid or collaboratory organisations.¹⁰

Here is another such passage:

Human resources are critical to getting cyberinfrastructure and applications working, keeping them working, and providing user support. In the interest of funding more grants, NSF has arguably under-supported the recurring costs of permanent staff, preferring to focus resources on direct research costs and ‘hard’ or ‘tangible’ assets. In the ACP [Advanced Cyberinfrastructure Program], human resources are the *primary* requirement in both development and operations, and success is clearly dependent on adequate funding both in centers and in the end-user research groups.¹¹

What is being said, in these passages, of the importance of discipline-specific computational support in the sciences could also be said of computational humanities or social science. And with respect to the nature of research projects themselves, it would be as true, these days, of the humanities as of science to say that ‘many contemporary projects require effective federation of both distributed resources (data and facilities) and distributed, multidisciplinary expertise, and that cyberinfrastructure is a key to making this possible.’

¹⁰ Daniel Atkins et al, *Revolutionizing Science and Engineering through Cyber-infrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure* (January 2003) 17 <http://www.communitytechnology.org/nsf_ci_report/>.

¹¹ Daniel Atkins et al, *Revolutionizing Science and Engineering through Cyber-infrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure* (January 2003) 60 <http://www.communitytechnology.org/nsf_ci_report/>.

What the Atkins report says—that one could *not* say of the humanities or perhaps even of social science—is that ‘prior investments provide a sound foundation for the ACP.’ In fact, prior investment in cyberinfrastructure for the humanities and social sciences is tiny, by comparison to what it has been in the natural sciences, and that puts us in a rather different position, with rather different needs. In addition to the two ‘resource constraints’ I discussed earlier—intellectual property and privacy rights—we have another resource constraint, and it is a constraint on a human resource, namely those disciplinary computational specialists. Education and training, therefore, must be even higher on the agenda for the humanities and social sciences than they are for computational science or computer science itself. Schools like the one I recently moved to, schools of library and information science, are our best bet for producing those specialists, I believe, and I firmly believe (with Margaret Hedstrom and John King) that libraries are one of the principal places that they will do their work. I also think, though, that we will have to reassert, more generally, the importance of mathematics in general education and in the liberal arts curriculum, beginning as early as middle school and high school. We will need English majors who have a background in logic, who can handle statistics, who do math, if we are going to turn out a generation of disciplinary specialists who can bring the accumulated wisdom of the humanities to bear in computational contexts—perhaps in helping build ontologies for scholarly projects in disciplinary contexts, or building tools for data-mining in the context of humanities research. I have met a few of them at the University of Virginia, and even graduated a couple with PhDs in English who have gone on to tenure-track appointments as humanities computing specialists in English departments; I am meeting and graduating a lot more of them now, at the University of Illinois’ Graduate School of Library and Information Science. These newly minted scholars, some of whom are specialists in disciplines of the humanities, or social sciences, and some of whom are specialists in information science, have arrived at that expertise without abandoning mathematics and logic. Consequently, they have absorbed and naturalised computational methods, and they hunger for more data. Given the necessary resources, they will—I am convinced—find novel ways to bring their disciplines to bear on the uncertainties, the

quandaries, the moral and aesthetic challenges, as well as the practical problems, of ‘the knowledge society.’

Those are some of my starting hypotheses, as chair of the ACLS commission on cyberinfrastructure for the humanities and social sciences. I expect that the work of the commission will test those hypotheses in various ways, and will alter them as a result. I present them here, not as predictions of the commission’s outcomes, but to mark my own starting point, and I would welcome your response to any of the points just raised. During the coming year [2005], the ACLS Commission is charged to:

- Describe and analyse the current state of humanities and social science cyberinfrastructure;
- Articulate the requirements and the potential contributions of the humanities and the social sciences in developing a cyberinfrastructure for information, teaching, and research; and
- Recommend areas of emphasis and coordination for the various agencies and institutions, public and private, which contribute to the development of this cyberinfrastructure.

CONCLUSION

I will close by observing that there is a kind of ten-year cycle to the sort of thing we are doing. Ten years ago, it was the National Information Infrastructure, and various commissions and committees around that term, that sparked a good deal of the priority-setting and decision-making that set the research agenda for the next decade. The humanities and the arts were a small part of that conversation, and there were some outcomes from that, but I hope this time around the engagement is a more profound one, and I hope the outcomes are more lasting. I also hope that the bridges we build in this process are bidirectional, and encourage collaborations and provocations that finally unite CP Snow’s two cultures, and deconstruct that binary opposition once and for all.¹²

¹² For further references, websites and initiatives in this area see: ACO*HUM (Advanced Computing in the Humanities, sponsored by the European Commission, published in 1999)

POSTSCRIPT

The report of the American Council of Learned Societies (ACLS) Commission on Cyberinfrastructure for the Humanities and Social Sciences titled '*Our Cultural Commonwealth*' was published in 2006. It proposed that 'an effective and trustworthy cyberinfrastructure for the humanities and social sciences will have the following characteristics:

- Be accessible as a public good
- Be sustainable

<<http://helmer.aksis.uib.no/AcoHum/book/>>; The National Information Infrastructure: Agenda for Action (1993) Ron Brown, Secretary of Commerce and Chair, Information Infrastructure Task Force <<http://www.ibiblio.org/nii/NII-Table-of-Contents.html>>; Humbul Humanities Hub <<http://www.humbul.ac.uk/>>; Voice of the Shuttle <<http://vos.ucsb.edu/>>; H-Net: Humanities and Social Sciences Online <<http://www.h-net.msu.edu/>>; NEH-funded online projects <<http://www.neh.gov/projects/online.html>>; Association for Computers and The Humanities <<http://www.ach.org/>>; Association for Literary and Linguistic Computing <<http://www.allc.org/>>; National Initiative for Networked Cultural Heritage <<http://www.ninch.org/>>; Consortium for Computers in the Humanities/Consortium pur ordinateurs en sciences humaines (coch-cosh) <<http://www.coch-cosh.ca/>>; Association for Computational Linguistics <<http://www.aclweb.org/>>; The American Association for History and Computing <<http://www.theahc.org/>>; Culture, Creativity and Information Technology, Social Sciences Research Council <<http://www.ssrc.org/programs/ccit/>>; History and Geography: Assessing the Role of Geographical Information in Historical Scholarship (2004) <<http://www.newberry.org/hgis/>>; Digital Resources in the Humanities <<http://www.drh.org.uk/>>; Inaugural Conference on Computational Social Science (2003) <<http://socialcomplexity.gmu.edu/5-2003conf/5-2003conf.htm>>; [and see The GIS History Project (1996), <http://www.geog.buffalo.edu/ngia/gishist/>, also Past Time, Past Place: GIS for History (ESRI: 2002)]; ACLS Occasional Paper No 41, *Computing and the Humanities: Summary of a Roundtable Meeting* (1998) <<http://www.acls.org/op41-toc.htm>>; ACLS Occasional Papers No 36, *New Connections for Scholars: The Changing Missions of a Learned Society in an Era of Digital Networks* (1997) <<http://www.acls.org/op36.htm>>; Institutional Models for Humanities Computing <<http://www.kcl.ac.uk/humanities/cch/allc/imhc/>>; Jahrbuch für Computerphilologie 4 (2002) <<http://www.computerphilologie.uni-muenchen.de/jahrbuch/jb4-content.html>>; *Humanities and Arts on the Information Highways* CNI/ACLS/Getty (1994) <<http://www.cni.org/projects/humartway/>>; Malhotra, Yogesh; Abdullah Al-Shehri and Jeff J Jones (1995) *National Information Infrastructure: Myths, Metaphors And Realities* <<http://www.brint.com/papers/nii/>>

- Provide interoperability
- Facilitate collaboration
- Support experimentation¹³

It further explained that ‘the necessary characteristics outlined above may be thought of as specifications for a humanities and social science cyberinfrastructure. Actually building something that answers to those specifications will require sustained effort and commitment in at least eight areas:

- Invest in cyberinfrastructure for the humanities and social sciences, as a matter of strategic priority
- Develop public and institutional policies that foster openness and access
- Promote cooperation between the public and private sectors
- Cultivate leadership in support of cyberinfrastructure from within the humanities and social sciences
- Encourage digital scholarship
- Establish national centers to support scholarship that contributes to and exploits cyberinfrastructure
- Develop and maintain open standards and robust tools
- Create extensive and reusable digital collections¹⁴

The report concluded by saying:

We should place the world's cultural heritage—its historical documentation, its literary and artistic achievements, its languages, beliefs, and practices—within the reach of every citizen. The value of building an infrastructure that gives all citizens access to the human record and the opportunity to

¹³ American Council of Learned Societies, *Our Cultural Commonwealth* (2006) <http://www.acls.org/cyberinfrastructure/OurCulturalCommonwealth.pdf> 27–29

¹⁴ American Council of Learned Societies, *Our Cultural Commonwealth* (2006) <http://www.acls.org/cyberinfrastructure/OurCulturalCommonwealth.pdf> 29–39

participate in its creation and use is enormous, exceeding even the significant investment that will be required to build that infrastructure. The Commission is also keenly aware that in order for the future to have a record of the present, we need legal and viable strategies for digital preservation; considerable investment is now required on that front as well. Investments need to be made on the basis of research, and, in this case, a good deal more research is needed on digital preservation, tools, and uses and users of digital collections, in academic settings and beyond.

But this is only part of the realization that the Commission hopes to leave with readers of this report. In a recent public presentation of the draft findings of this report, the Commission's chair was asked, "If your report were a complete success, what would be the result, five or six years from now?" The answer is twofold. First, if this report's recommendations are implemented, then in five or six years, there will be a significantly expanded audience for humanities and social science research among the general public. A relatively small audience on the open Web will still be a far larger audience than scholars in these disciplines have been able to find up to now in academic bookstores, research libraries, and print journals. Second, if the recommendations of this report are implemented, humanities and social science researchers five or six years from now will be answering questions that today they might not even consider asking.¹⁵ (footnotes omitted)

¹⁵ American Council of Learned Societies, *Our Cultural Commonwealth* (2006)
<http://www.acls.org/cyberinfrastructure/OurCulturalCommonwealth.pdf> 40

PART TWO

INSTITUTIONAL AND LEGAL
FRAMEWORKS

DESIGNING INSTITUTIONAL INFRASTRUCTURES FOR e-SCIENCE *

Paul A David¹ and Dr Michael Spence²

1. INTRODUCTION – THE OPPORTUNITIES AND CHALLENGES OF e-SCIENCE

1.1 Background, Motivation and Purpose

The opportunity exists today for unprecedented connections between scientists, information, data, computational services, and instruments

* This work draws extensively on the text of the authors' research memorandum 'Toward Institutional Infrastructures for e-Science: the Scope of the Challenge', released in May 2004 as *Oxford Internet Institute Research Report No 2* (hereafter 'David and Spence'). That publication was derived from the OII's (September 15, 2003) Final Report on the 'The Institutional Infrastructure of e-Science: The Scope of the Issues,' a project undertaken with the support of a grant to the University of Oxford from the Joint Information Systems Committee (JISC) of the UK Research Councils. It does not include material from the *Research Report No 2* Addenda (Appendixes 1–7), references to which appear in the text and footnotes notes. Interested readers can access and print the appendices from pages 63–89 of the e-document linked to the following webpage: <<http://www.oii.ox.ac.uk/research/project.cfm?id=26>>. The contributions made to the work of the original project by many people and organisations have been acknowledged with gratitude on earlier occasions. But it is nonetheless appropriate here to recall the excellent research assistance on the JISC Project that was provided by Matthijs den Besten (now at the Oxford e-Research Centre), and the insightful and constructive comments on earlier drafts by W Edward Steinmueller (University of Sussex) and Paul Uhler (National Academy of Sciences).

The present text also incorporates material drawn from a related paper by P A David, 'Towards a cyberinfrastructure for enhanced scientific collaboration: providing its 'soft' foundations may be the hardest part', SIEPR Policy Paper 04–001 (May 2004), Stanford University (a May 2005 revision is available at <<http://siepr.stanford.edu/papers/pdf/04-01.html>>). David acknowledges the continuing support that his research has received from the Stanford Institute for Economic Policy Research and the Oxford Internet Institute.

¹ Professor of Economics, Stanford University; Emeritus Fellow of All Souls College, in the University of Oxford and Senior Fellow of the Oxford Internet Institute.

² Vice Chancellor, University of Sydney, formerly Head of Social Sciences Division, the University of Oxford, Fellow of St Catherine's College and Faculty Associate of the Oxford Internet Institute.

through the Internet. A new generation of information and communication infrastructures, including advanced Internet computing and Grid technologies, is beginning to enable much greater direct and shared access to more widely distributed computing resources than previously has been possible.³ The term ‘e-Science’ usually is applied in reference to large scale science that, increasingly, is being carried out through distributed global collaborations enabled by the Internet.⁴ Such collaborative scientific enterprises typically require access to very extensive data collections, very large scale computing resources, and high performance visualisation of research data and analysis of results by the individual users. The potential for these advances in technology to support new levels of collaborative activity in scientific and engineering, and ultimately in other domains, is a major driving force behind the UK’s Core e-Science Programme.⁵

A growing number of those acquainted with these technological developments anticipate that they will have transformative effects on the organisation and conduct of ‘knowledge work’ – particularly scientific and engineering research. Thus, the 2003 report by a distinguished

³ General overviews of the Grid and related Internet computing are provided by I Foster, ‘Internet Computing and the Emerging Grid’ (2000) 7 December *Nature* <<http://www.nature.com/nature/webmatters/grid/grid/html>>; I Foster, ‘The Grid: Computing without Bounds’ (2003) April *Scientific American*. For further detail, consult I Foster, I Kesselman and C Kesselman (eds), *The Grid: Blueprint for a New Computing Infrastructure* (2001); I Foster et al, *The Physiology of the Grid* (Version 2/17/2002) <<http://www.globus.org/research/papers/ogsa.pdf>>.

⁴ See David and Spence (2003/2004: Appendix 1 Computer-mediated telecommunication network supports for collaborative research activities: concepts and definitions) on this and related terminology found in the text. For an overview of connections between the UK e-Science Programme, Grid services and high bandwidth middleware, by the e-Science Core Programme’s Director, see: T Hey, *Towards an e-Science Roadmap*, <http://umbriel.dcs.gla.ac.uk/nesc/general/news/ukroadmap180402/TonyHeyTowards_an_eScience_Roadmap.pdf>.

⁵ In November 2000 Dr John Taylor (the Director General of the Research Councils) announced £98M funding for a new UK e-Science programme: £15M was allocated by the Office of Science and Technology (OST) to the Core e-Science Programme, a cross-Council activity to develop and broker generic technology solutions and generic middleware to enable e-Science and form the basis for new commercial e-business software. OST funding for the core e-Science Programme has been augmented by a further £20M from the Department of Trade and Industry (DTI), which is to be matched by £15M from industry. See UK Research Council e-Science Programme <<http://www.research-councils.ac.uk/escience/>>; also, for allocations to specific science domains <<http://www.escience-grid.org.uk/docs/briefing/funding.htm>>.

advisory panel to the NSF Directorate of Computer and Information System Engineering, chaired by Daniel Atkins, envisaged an enhanced computer and network technologies supporting those connections as forming a vital infrastructure – dubbed the *cyberinfrastructure* – whose effects would be analogous to the historical impacts of super-highways, electric power grids, and other physical infrastructures in raising the productivity of conventional work.⁶ The recommendations of the ‘Atkins Committee Report’ were swiftly embraced by the NSF, which established a high level Office of Cyberinfrastructure and in 2005 tasked multi-disciplinary, cross-foundational teams to further elaborate a ‘vision’ that would guide the Foundation’s program of cyberinfrastructure (CI) investments in four overlapping and complementary areas. These were 1) High Performance Computing, 2) Data, Data Analysis, and Visualisation, 3) Cyber Services and Virtual Organisations, and 4) Learning and Workforce Development. Following an extensive process of consultation on drafts, a comprehensive and no doubt influential ‘vision statement’ has received endorsement from the newly constituted Cyberinfrastructure Council (CIC) within the Foundation.⁷

The original central expectation animating this initiative was that the solution of *technical* problems associated with an advanced cyberinfrastructure would unleash new scientific capabilities – leading to key discoveries, such as improved drug designs, deeper understanding of fundamental physical principles, and more detailed environmental models. With the passage of time has come explicit recognition that in reality, such gains, if they materialise would likely be the *combined effect of social and technical* transformations. Indeed, the prefatory letter from NSF

⁶ The potential to revolutionise science and engineering in the 21st century is set out at some length as the rationale for a major programmatic commitment by NSF, in D E Atkins et al, *Revolutionizing science and engineering through cyberinfrastructure: Report of the National Science Foundation blue-ribbon advisory panel on cyberinfrastructure* (February 2003)

<http://www.communitytechnology.org/nsf_ci_report/>. On the transformative implication in the local, Oxford context, see also, P Jeffries, ‘e-Science and the Grid: Why it will change Oxford’, (Presentation by the Director of the Oxford University e-Science Centre to the Oxford BioInformatics Forum, 7 November 2001) <<http://e-science.ox.ac.uk/>>.

⁷ See National Science Foundation, *Cyberinfrastructure Vision for 21st Century Discovery* (March 2007) <http://www.nsf.gov/od/oci/CI_Vision_March07.pdf>.

Director Arden L Bement, introducing the *Cyberinfrastructure Vision for 21st Century Discovery*⁸ speaks less of network engineering than of cultures:

At the heart of the cyberinfrastructure vision is the development of a cultural community that supports peer-to-peer collaboration and new modes of education based upon broad and open access to leadership computing; data and information resources; online instruments and observatories; and visualization and collaboration services. Cyberinfrastructure enables distributed knowledge communities that collaborate and communicate across disciplines, distances and cultures. These research and education communities extend beyond traditional brick-and-mortar facilities, becoming virtual organizations that transcend geographic and institutional boundaries.

Yet, as one pursues the specifics of the cyberstructure vision, even in the chapter on ‘virtual organisations’, the original conceptualisation resurfaces; the over-riding theme is that these social formations will be called forth more or less automatically by the empowering features of the new collaboration technologies and data resources that are being promised to geographically distributed researchers. Here is one of the less technically detailed, but nonetheless emblematic expressions of that faith:⁹

The convergence of information, grid, and networking technologies with contemporary communications now enables science and engineering communities to pursue their research and learning goals in real-time and without regard to geography ... the creation of end-to-end cyberinfrastructure systems – comprehensive networked resources – by groups of individuals with common interests is permitting the establishment of Virtual Organizations (VOs) that are revolutionizing the conduct of science and engineering research and education. A VO is created by a group of individuals whose members and resources may be dispersed geographically and/or temporally, yet who function as a coherent unit through the use of end-to-end cyberinfra-

⁸ (2007) i.

⁹ *CI Vision* (2007) 32.

structure systems. These CI systems provide shared access to centralized or distributed resources and services, often in real-time. Such virtual organizations supporting distributed communities go by numerous names: collaboratory, co-laboratory, grid community, science gateway, science portal, and others.

One might be reasonably confident about the pace and scope of future technical advances in computing that will follow from the dynamics of ‘Moore’s law’, the plummeting price-performance ratio of micro-processors, and an enormous expansion of digital bandwidth and inexpensive memory. But, whereas as far greater uncertainties continue to surround the extent to which individuals, groups, organisations, and institutions will be able adapt to and benefit from the novel technological systems that may be engineered on those foundations, even enthusiastic advocates of heavy investment in hardware, middleware and software components of the coming ubiquitously accessible computational facilities of a Grid-services enabled cyberinfrastructure, have come to recognise that there may be a profound gap between ‘raw’ performance capabilities (based on bandwidth, storage capacity, processor speed, and interconnection protocols), and its realised performance (based upon the ‘usability’ properties of the constituent system designs). Some of that awareness has been heightened by recalling the disappointing findings of systematic evaluations of the pioneering *collaboratory* projects – which had been mounted in the US during the early 1990’s to explore the potentialities of the ‘virtual laboratory’ concept.¹⁰ That, however, is not the only gap that can significantly limit the transformational potential of ‘cyber-infrastructure investments’.

Achieving the aims and aspirations of e-Science and the Cyberinfrastructure vision is not just a matter of breakthroughs in hardware or software engineering, or system design improvements to provide tools that will be readily useable by individual researchers and their organisations – as challenging as those engineering tasks may be, or

¹⁰ See, for example, Thomas A Finholt, ‘Collaboratories as a new form of scientific organization’ (2003) 12 (January) *Economics of Innovation and New Technology* 5–25. David and Spence (2003/2004: Appendix 1, and 2, below) present some information about the characteristics of the pioneer ‘collaboratories’.

even the development of programs to train researchers and teachers in the effective use of the new tool set. The informal norms and formal rule structures for collaboration on the ground as well as in cyberspace, that is to say, the ‘institutional’ contexts within which the work of communities of scientific and technical researchers is carried on, also will matter profoundly. These, too, will constrain as well as facilitate improvement in the effectiveness of the variety of research collaborations that actually are formed within and across disciplinary, university, and national boundaries.

The institutional and organisational ‘environment’ of public sector e-Science encompasses a wide and diverse array of interrelated social, economic and legal factors that shape the utilisation, consumption, governance and production of e-Science capabilities and artefacts. Principal amongst these are the following three:

- the rules and regulations of the agencies that provide grant and contract funds to researchers in public research organisations;
- the latter organisations’ own rules and administrative procedures governing formal relationships with their employed research staff (and research students, in the case of universities), which typically will refer to elements of the external legal system (such as the statutes governing contracts, liability, privacy and intellectual property);
- informal epistemic community norms and conventions, which will be recognised if not always adhered to by members of the various scientific and technological professional groupings, as well as some particular ‘local social norms’ that are likely to emerge among colleagues engaged in extended research projects.

Thus, any systematic approach to the transformation of the conduct of scientific and technological research hardly can avoid directing attention to these ‘institutional infrastructures’; their features are likely to turn out to be quite crucial for ensuring that the technical capabilities of advanced Internet computing and the Grid actually will be accessed, effectively

applied and exploited thoroughly by researchers organising collaborations in a variety of fields. In Figure 1 (below) the foregoing non-technological elements are depicted, along with the middleware platforms and supporting layer of computer mediated communications hardware and software, as providing key infrastructural and regulatory supports of the 'e-Science collaboration domain'. It will be noticed that each the four 'facets' of the tetrahedron in Figure 1 makes contact with, and hence is both bounded and supported by three other elements of the 'infrastructure'. None of the elements exists in isolation, and hence in the long run it is appropriate to view all of them as endogenously.¹¹

In shifting the focus of attention from questions of technical engineering to institutional design, is it particularly important to bear in mind that the goals and requirements of the research organisations and host institutions that are likely to emerge as the eventual users of these facilities may well diverge significantly from those found among the projects which today are pioneering the development of hardware and software systems for e-Science. Some forward-looking exercise of the imagination, therefore, is in order at this time, contemplating the likelihood of e-Science collaborations that will not bear close resemblance to the projects that currently are proceeding under that banner.

To hope to avoid, or even to significantly postpone the effort of critically thinking through the likely needs of projects that have yet to be conceived of, may well prove be a costly strategy. Very substantial resource costs can be entailed when societies try to utilise technological systems the immediate applicability of which turn out to be unexpectedly limited outside the immediate conditions of their genesis. Figure 1 The e-Science Collaboration Domain and Its Infrastructural & Regulatory Supports is a concrete, pertinent and not unfamiliar illustration to support that proposition is available in the story of the evolution of the ARPANET into the Internet.¹² The ARPANET and its direct

¹¹ That general perspective informs the approach taken by this report, but is not explicitly elaborated. For further discussion, see P A David, D Foray and J Mairesse, *Public dimensions of the knowledge-driven economy: an analytical framework*, (21 June 2001) Working Group on the Knowledge Economy, Center for Education Research and Innovation (CERI), OECD.

¹² This draws (briefly) on P A David, 'The Evolving Accidental Information Super-Highway' (Fall 2001) 17(2) *Oxford Review of Economic Policy* (Special Issue: *The Economics of the Internet*) 159–87.

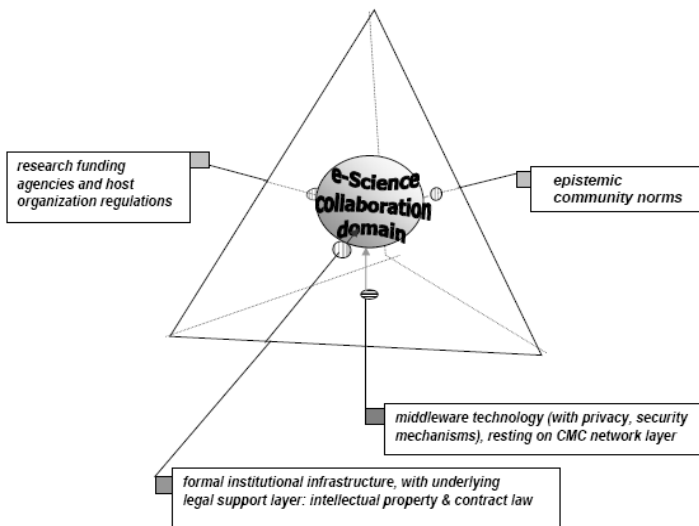
successor, the NSFNet, were communications infrastructures that had been developed under public auspices to serve the needs and circumstances of university-based research groups. This was an environment of application in which individual and organisational behaviours generally are regulated tightly by social norms and institutional rules, and where the dominant ethos is that of co-operation in non-commercially oriented activities of inquiry and information dissemination. The influence exerted over the course of three decades of development work (1964–1994) by the social parameters of that essentially stable background is reflected in the technical specifications of the Internet's end-to-end architecture and the TCP/IP protocol stack. Rather unexpectedly, however, the rapid privatisation and commercialisation of the new, 'connection-less' communications facility that took place during the mid-1990s had the effect of transferring this technological artefact into a social environment that was very different from the one in which it had been designed. The consequences have not been entirely unproblematic, to say the least.

Although the Internet has now begun to be used extensively for commercial purposes, this new context of use is one for which the network of networks has been revealed to be less-than-optimally suited, in more than one respect. This has given rise to an important challenges in areas such as: finding appropriate business models for the Internet's open architecture and culture; adjusting the intellectual property rights (IPR) regime to the new structure of information reproduction and transmission costs; filtering unwanted (spam!) messages; providing security and protection from malicious, or simply non-co-operative, actions by other system users; and designing quality of service (QoS) suitable for network services that was never contemplated in the original system design.

Early experience with the Internet and collaboration-supporting technologies suggests that data and other resource sharing across the institutional boundaries of the resulting 'collaboratories' – which are at the heart of e-Science – already is confronting legal and political administrative constraints, particularly those arising around intellectual property rights issues. Experience in other institutional settings, such as

e-government, reinforces these observations.¹³ Finding even reasonably satisfactory solutions to such academic and commercial challenges encountered by the first generations of Internet users undoubtedly has created many opportunities for ingenuity to display itself, including some quite profitable new lines of business. Yet, all this adaptive effort, whether successful or not, has entailed considerable unanticipated costs. The open question is whether we are now in a position to make better preparations to utilise the new e-Science tools that it appears to be within our power to fashion.

Figure 1



¹³ For example, e-Government research underway at the University of Southern California's Information Systems Institute (ISI) has focused on the development of middleware to support information sharing among government agencies. Technical advances have been demonstrable, but entrenched political administrative traditions in US federal agencies, which are quite different from academia, have limited the ISI's ability to work with data of central importance, such as the US Census.

Motivated by the foregoing contemporary developments, and informed by the historical experience of the Internet's origins in the work of university-based scientists and engineers, this essay has a three-fold purpose:

First, to articulate the nature and significance of the non-technological issues that will bear on the practical effectiveness of the hardware and software infrastructures that are being created to enable collaborations in *e-Science*;

Second, to succinctly characterise the fundamental sources of the organisational and institutional challenges that need to be addressed in regard to defining terms, rights and responsibilities of the collaborating parties, and to illustrate these by reference to the limited experience gained to date in regard to intellectual property, liability, privacy, and security and competition policy issues affecting scientific research organisations;

Third, to propose approaches for arriving at institutional mechanisms whose establishment would generate workable, specific arrangements facilitating collaboration in *e-Science*; and, that also might serve to meet similar needs in other spheres such as *e-Learning*, *e-Government*, *e-Commerce*, *e-Healthcare*.

1.2 Organisation and Overview

The main body of this chapter is organised in four main parts. The following sections of Part 1 address the first of the three principal tasks that have just been described, beginning with an examination of the technological and institutional contexts of *e-Science* and their interrelated dynamic evolution (in section 1.3), and then reviewing the different ways in which these two infrastructure components affect the costs of collaboration (in section 1.4).

The second principal task is the subject of Part 2, which opens by considering the social conventions and norms that may be said to govern collaboration within scientific collaborative workgroups,

distinguishing these from the agreements governing the contractual relations among the institutions in which the members of those work-group are employed. Section 2.1 takes notice of the additional elements complexities that are presented by recent changes in academic communities' professional norms and in the career conditions affecting university researchers. Still further complications in the relationship between informal and formal governance mechanisms are seen to arise as a result of the ways in which uncodified normative structures governing professional conduct among collaborating individuals may be affected by the introduction of explicit legal contracts with their employing institutions, as well as with their academic peers.

Section 2.2 takes two necessary analytical steps towards understanding the character of the challenges involved in finding or designing new concrete institutional arrangements that facilitate the formation and conduct of socially productive research collaborations in e-Science. Those challenges are rendered more formidable by the need to provide for multi-institutional collaborations including international partners, as well as to accommodate trans-disciplinary projects involving distinct research units (departments, laboratories, institutes or centres) within the same Public Research Organisation (PRO).

The first step is to delineate (in sub-section 2.2.1) the variety of distinct benefits offered by co-operative organisation of research, and the multiple sources of conflict that are likely to exist among the interests of the potential collaborators, as well as among the administrative entities and employing institutions. The second step (in sub-section 2.2.2) draws upon recent economic analysis to suggest reasons why the social efficiency of publicly funded research investments is likely to be greater under a regime of more liberal contractual arrangements among the individual collaborating researchers; and, particularly, under rules that provide 'weak' rather than 'strong' protection for the commercial exploitation of intellectual property rights held individually by the participating institutions. From this analysis there follows an important meta-principle: the appropriate approaches to the institutional design problem for publicly funded collaborative e-Science are those that would be especially responsive at the margin to satisfying the collaborating agents' organisational requirements for conducting the research in question. Correspondingly, they will be less disposed to accommodate

other incompatible corporate interests and goals that their respective administrative units and employing institutions may seek to attain through participation in the proposed collaboration.

Section 2.3 considers four classes of legal problem that might arise in the context of collaborative e-Science. There has been considerable discussion of the impact of intellectual property rights on scientific collaboration. Getting the balance wrong between the ownership of, and access to, knowledge resources entails serious social costs that recently have been perceived more widely beyond the boundaries of the scientific community. But, it is surprising how few people have recognised that intellectual property rights are only one among the many kinds of legal issues that need to be successfully resolved to facilitate collaborative work. Collaboration among researchers can be affected by the entire complex of legal norms and informal professional conventions. It is important that institutional arrangements are made so as to minimise the extent to which the law becomes an impediment to cooperation among researchers, whether directly, or indirectly by undermining informal mechanisms of trust and dispute resolution.

Four different types of legal problem that a collaborative project might encounter are examined *ad seriatim* (by sub-sections 2.3.1 through 2.3.4, respectively). These problem-classes are concerned with:

1. the legal relationships among the parties to an e-Science collaboration, particularly where some of the parties are operating in different jurisdictions;
2. the materials that each party brings to a collaboration;
3. the resources, if any, to which the collaborative project will give rise;
4. the apportionment (among the parties) of liability for potential harms arising from the collaborative project.

The third of the major tasks is taken up in Part 3. An initial assessment is made (in section 3.1) of some oft-recommended legal approaches to simplifying institutional mechanism design problems, notably by introducing standard form agreements, and by harmonising disparate and potentially conflicting legal doctrines, and statutes. Finding many practical deficiencies in the latter approaches, the argument (in section

3.2) favours developing an alternative, more flexible, modular process to generate contextually appropriate contractual arrangements for collaborative research projects. The proposed process calls for the principal public funding bodies engaged in building the technological components of the e-Science infrastructure also to lead their authority to the work of a new ‘public actor’ in the shape of an independent advisory body on institutional infrastructures. The role of that body, referred to as the Advisory Board on Collaboration Agreements (the ‘ABCA’) in e-Science, would be to guide and co-ordinate the formulation and dissemination of an array of specific contractual clauses that could be used to construct a variety of legal agreements governing scientific and technological research collaborations among universities and other corporate partners. The ABCA also could develop sufficient expertise to provide guidance for research groups seeking effective informal arrangements to deal with various internal governance issues, thereby facilitating the more spontaneous, ‘bottom up’ formation of projects enabled by the emerging e-Science collaboration tools.

The remaining sections of Part 3 set out a number of requirements that should be met by the constitution of such an advisory body, in terms of the private expertise and public agency experience and interests upon which it would need to be able to draw (section 3.3), and the development of an evolving informational base about the actual collaboration arrangements and their efficacy (section 3.4).

In the fourth and concluding Part, two different questions are treated. They are related, however, inasmuch as each has a concrete bearing upon the practical implementation of the general approach, and the specific recommendations advanced by Part 3. The first of the pair concerns the nature of the broad principles that the proposed ABCA would embrace and seek to embody in a menu of contractual clauses; whereas the second addresses the need to find an expedient ‘way forward’ that would provide near-term guidance for issues of governance arising from the e-Science projects that are presently underway.

Part 3’s procedural recommendations for ‘contractually constructing’ arrangements to support a variety of e-Science collaborations (intentionally) are formulated in a way that remains neutral with respect to the general thrust of the guiding principles that the ABCA would

embrace. Part 4, by contrast, takes notice of the growing number of calls for modifications in the intellectual property rights policies of governments – particularly in order to assure more protection of the public domain in scientific and technical data and information. Policy statements in that vein have emerged recently from a number of influential bodies in the United Kingdom, the European Union and the United States. It would be remarkable were the ABCA not to give weight to these concerns in delineating the principles against which it should assess proposed model contractual clauses for use in e-Science collaboration agreements.

Indeed, it is a positive advantage of the approach based upon contractual agreements that it would allow a direct and positive response to the worries expressed about excessive restraints being imposed upon open scientific collaboration by excessive recourse to intellectual property rights (IPR) protections. In other words, there is a case (developed in section 4.1) for using the establishment of the ABCA as a means to avoid having to wait for statutory reforms in the IPR regime. As an illustrative case in point, it is shown how public licensing of intellectual property under a standard form of ‘free and open source software’ (F/OSS) license—such as the GNU General Public License—may be used in conjunction with private contractual provisions governing commercial exploitation of the ‘open source’ code. This ‘dual licensing’ approach can accomplish two purposes that often are presented as inimical from a practical standpoint. On the one hand, public funding authorities may regard it as efficient to maintain the essential functional features of a ‘knowledge commons’ in the software tools and products whose creation they have sponsored, but, at the same time, it may be desirable to leave some scope for market incentives to mobilise complementary private sector investment directed toward further development of basic software innovations released under the terms of F/OSS licenses.

The solution suggested is to allow ‘dual licensing’ of some categories of publicly funded software (and middleware), combining GNU GPL licenses with the option of constructing contractual arrangements (built from standard clauses) whereby third parties obtain the copyright holder’s permission to develop modifications and extensions for private commercial distribution. Under this approach, there would be a clearly identifiable need for the services of an advisory board-like entity to

develop appropriate contractual clauses that would work in conjunction with copyright licenses that embody the so-called ‘copyleft’ principle. To initiate effective implementation of the recommended contractual customisation approach, and its application in the context of ‘dual licensing’ of F/OSS products, will require a strongly supportive stance on the part of the national and international public agencies and private foundations that provide major funding for the e-Science programs and projects conducted in PROs.

The discussion of practical measures closes (in section 4.2) by outlining an interim course of action for the e-Science Core Program to follow, in order to furnish itself with expert advice and counsel for the decisions that must be made about non-technological governance issues affecting the use of the software systems whose creation it has sponsored, and expects to sponsor in the foreseeable future. This ‘way forward’ could have a potent impact, not only in shaping the near-term institutional environment for e-Science in the UK, but by initiating the first steps on a transition path towards the eventual institutionalisation of an independent ABCA along the lines envisaged here. By moving quickly to establish an interim Working Party on institutional infrastructures for e-Science, and having that body actively engage with representatives from international counterpart programs, the UK’s Core Program soon could begin exerting significant international influence. It would thus move closer to fulfilling the promise of e-Science to accelerate advancement of knowledge and material well-being on a global scale.

1.3 Technological and Institutional contexts of e-Science

e-Science is a term used increasingly widely as a generic label for all scientific and technical research activities conducted on the Internet.¹⁴ But it is employed more specifically here, in referring to scientific activities supported by high bandwidth computer-mediated telecommunications networks, and particularly to encompass the variety of such digital information-processing applications that are expected to be enabled by the *Grid*. The latter may be viewed as the general purpose

¹⁴ See David and Spence (2003/2004: Appendix 1) for a glossary of descriptive terms in the text, including: e-Science, the Grid, Collaboratories, Virtual Laboratories, Cyberinfrastructure, and their relationship to one another.

network technology which will serve to facilitate new, computationally intensive forms of scientific inquiry: desktop supercomputing, distributed supercomputing (a marriage of parallel and distributed computing), extensive exploration of linked distributed dynamic databases by high-speed search engines, and collaborative environments (collaboratories or virtual laboratories) including smart instruments for data capture and analysis that are coupled to supercomputing resources, and so on.

Collaborative e-Science is the aspect of the vision of 21st century science that holds out the most exciting new possibilities, and which also poses the most demanding challenges at the technical, social and legal levels. Technological and social changes are intertwined, and in many respects their interactions and mutual adaptations are difficult to disentangle. It is undoubtedly the case that technological advances that have placed new, more productive and more costly facilities and instruments at the disposal of researchers are prominent among the forces driving the widely observed trend towards collaborative organisation of scientific inquiries. But, beyond the increasing scale of projects utilising 'lumpy' capital-intensive facilities in fields such as physics and astronomy, the sheer increase in the amount of pertinent information, and the progressively more specialised knowledge and expertise that must be brought to bear in order to conduct fruitful research programs in most branches of science, have contributed to the growth in the size of teams and the numbers of co-authors on scientific, technical and scholarly publications.¹⁵ Although the continuous pressure toward specialisation and division of labour has pushed researchers into the forms of cooperative knowledge transactions entailed in collaborative inquiry, more recently researchers located at widely dispersed institutions have been drawn into informal and formal collaborations by the dramatic

¹⁵ Although the emergence of research collaboration in 'Big Science' fields was viewed from the 1960s onwards as a significant novelty reflecting underlying tendencies in the organisational structure of modern science, the increasing generality of collaborative organisation is now attracting fresh interest as the most recent phase in a broader, longer and more continuous development. See, for example, J S Katz, 'Geographical proximity and scientific collaboration' (1994) 31(1) *Scientometrics* 31–43; D Hicks and S J Katz, 'Science policy for a highly collaborative science system' (1996) 23(1) *Science and Public Policy* 39–44; J S Katz and B R Martin, 'What is research collaboration?' (1997) 26(1) *Research Policy* 1–18; H Etzkowitz and C Kemelgor, 'The Role of Research Centres in the Collectivization of Academic Science' (1998) 36 *Minerva* 271–88.

advances that have been achieved in computer-mediated telecommunications.¹⁶

However difficult it may be to empirically identify the separate influences of the technological from those of the other, social factors affecting collaborative research, it is not only possible but also important to draw this conceptual distinction. This chapter goes further, however, by directing special attention to the legal-institutional contexts of the array of collaborative research activities that are expected to be greatly facilitated by improvements of the technological components of the e-Science infrastructure. As a background for the discussion, a taxonomic framework for e-Science collaborations has been developed that highlights the various classes of interactions among collaborating parties that these technical facilities can support. This framework classifies collaborations on the basis of their major purpose, rather by reference to the particular digital information tools and services they might employ. Our taxonomy distinguishes among virtual laboratory activities conducted via the (enhanced) Internet that are pre-dominantly:

1. 'community-centric' – aiming to bring researchers together either for synchronous or asynchronous information exchanges;
2. 'data-centric' – providing accessible stores of data captured or extracted from remote sources, and creating new information by editing and annotating them;
3. 'computation-centric' –providing high-performance computing capabilities either by means of servers accessing super-computers and parallel computing clusters, or making possible for the collaborators to

¹⁶ See David and Spence (2003/2004: Appendix 1, Figures 1–3). This is reflected in the rising frequency of inter-institutional collaborative publications among US university researchers in scientific domains where average team size is comparatively small (for example, mathematics, and economics); and also by the observation that the growth of inter-institutional collaborative publications involving US academic researchers in international teams has outpaced that of purely domestic inter-institutional collaborative publications, as well as the rate of growth in average team sizes. See James D Adams et al, *Patterns of Research Collaboration in U.S. Universities, 1981–1999*, (March 2002) Economics Department Working Paper, University of Florida.

- organise peer-to-peer sharing of distributed computation capacity;
4. 'interaction-centric – enabling applications that involve real-time interactions among two or more participants, for decision-making, visualisation or continuous control of instruments.

When this scheme is applied to classify the array of Pilot Projects that have been funded under the e-Science Core Programme in the United Kingdom, the data-centric branch of the taxonomic tree emerges as far and away the most densely populated.¹⁷ The situation contrasts with the more uniform distribution that emerges from a comparable classification of much small number of pioneer collaboratory projects that were organised under public funding programs in the US during the late 1980's and early 1990's.¹⁸ That difference reflects in part the focus of the e-Science program on the creation of middle-ware platforms and tools, and in part the greater centrality of the roles that digital databases have more recent come to occupy in the work of science and engineering communities. Yet, a suspicion remains that some influence on the profile of these the Pilot Project sample also has been exerted by consideration of the greater administrative complexities that would have to be overcome to organise more thoroughly interactive modes of collaboration among research groups situated at various institutions within the UK.

The institutional infrastructure for e-Science collaborations might be viewed by some to be the soft part of what the report of the recent NSF Blue Ribbon Advisory Panel¹⁹ refers to as the 'Cyberinfrastructure' that promises to revolutionise science and engineering.²⁰ But, in truth, its design, construction, maintenance and updating pose many challenges

¹⁷ Another use of the taxonomy has been to assist in identifying a subset among the Pilot Projects that contained representatives of each of the 'collaboration purposes', and whose activities could be studied more closely in order to understand the variety of e-Science research contexts for which supportive institutional arrangements would need to be constructed.

¹⁸ See David and Spence (2003/2004: Appendix 1, Figure 4).

¹⁹ (February 2003).

²⁰ See Atkins et al, *Revolutionizing science and engineering through cyberinfrastructure* (2003) <http://www.communitytechnology.org/nsf_ci_report/>.

that are at once more delicate and harder than the technical engineering feats required for reliable and secure Grid-enabled computing. Institutions simultaneously are run by and govern human agents, and, for that reason as well as others, they are considerably less plastic than most machine organisations – i.e., systems composed of technological artefacts. Often, when they function well, institutions and the behavioural norms they reinforce become unobtrusive and tend to disappear into the background, so that the question of whether they will require modification to continue functioning smoothly in new environments is often deferred until after those environments have materialised.

Scientific teams engaged in hardware and software engineering in order to forge the tools needed to support their own work are, as a rule, more than fully tasked. They seldom are able to focus concurrently on the issues of how social and technological mechanisms can best be combined to address the array of complex problems that other users of those tools eventually would need to solve before the potentialities e-Science can be fully realised. Nor should working scientists be expected to possess the necessary expertise to consider the problems of developing procedural norms and formal contractual arrangements governing collaborative contributions of research resources.²¹

1.4 Towards Envisaging the Cyberinfrastructure with Collaboration Costs

The functional domain of institutional arrangements supporting scientific collaboration is both extensive and complex. These arrangements will govern the terms of access to and control over instruments and other physical facilities, and the data-streams generated in the research process. They will, in effect, apportion the scientific recognition and the disposition of ownership rights in collective work products created in cyberspace. They must also assign responsibilities for errors of commission and omission in those research outcomes, as

²¹ Lest there be any doubt on this score, it should be emphasised that the limited attention accorded to institutional design by scientists and technologists is a consequence of specialisation, and therefore the comment here is not meant as a reproach. As Bertrand Russell said, ‘We forgive specialists, because they do good work.’

well as liabilities for damages and legal infractions of various kinds arising from the actions of participants in the joint activities.

Generic collaborative arrangements of these kinds involve issues whose solutions naturally may appear quite familiar, and altogether tractable in the context of a co-located research team. Yet, the same issues quickly can become dauntingly complex when collaboration is extended to a multiplicity of geographically distributed teams and physical facilities, each of whose members have contractual relationships as employees of, or consultants to one or another among several different corporate entities. The latter, moreover, may well mix both public and private sector institutions and organisations all of which are not situated within and hence under the governance of a single legal jurisdiction and political authority.

It is evident that the complex collaborative undertakings in view here – those that are meant to be enabled, indeed, empowered by e-Science facilities and services – cannot be supposed to arise and function automatically as ‘perfect teams’ expressing some primitive cooperative impulse among the human actors. Quite the contrary: the collaborators will need to find solutions for non-technological issues of resource allocation and governance that involve conflicts arising from the divergent interests of the individuals and organisations involved. Moreover, to sustain extended programs of research that continue to build upon and utilise the specialised knowledge that they generate, those solutions must be sufficiently flexible to accommodate the high order of uncertainty that inevitably surrounds research activities. That is especially so for fundamental, exploratory research programs of the sort for which public support is particularly warranted. Only the satisfactory resolution of those conflicts will permit realisation of the gains from cooperation. Yet, it is important not to lose sight of the reality that ‘conflict resolution’ is not a costless process. Consequently, the means by which such solutions are arrived at ought not impose heavy ‘transactions costs’ upon the parties, thereby draining resources from the conduct of research itself, or, worse still, undermining whatever cooperative spirit and ethos of common purpose initially animated the collaborative enterprise.

The lattermost of these requirements may be seen to be present in the very idea of an e-Science ‘infrastructure’ as that is now coming to be

conceptualised. The recent report of the NSF Blue Ribbon Advisory Panel on Cyberinfrastructure²² describes the latter concept in expansive terms, in which the activities of human agents and organisations also are subsumed under the heading of infrastructure. According to the Panel,²³ whereas, historically ‘infrastructure’ has been viewed by people in the computer and telecommunication engineering sciences ‘largely as raw resources like compute cycles or communication bandwidth,’ now it is critical to think of [cyber-] ‘infrastructure’ as having three rather different basic components:

- **Technological artefacts.** These human-constructed artefacts include facilities (computers, mass storage, networks, etc) and software. These artefacts sometimes provide services, and sometimes they are simply available to be ‘designed into’ applications.
- **Technological services.** Various capabilities are provided as services available over the network rather than as software artefacts to be deployed and operated locally to the end-user.
- **Services from people and organisations.** These include everybody who is providing a shared pool of expertise leveraged by the entire scientific and engineering research community to develop and operate the technological artefacts and provide advice and assistance to end-users making use of them.’

Given the inclusion of the lattermost among these, it is rather remarkable that nothing in the report of this NSF Panel addresses the nature of the institutional settings, the incentive mechanisms, and the organisational culture of those who are ‘providing a shared pool of expertise leveraged by the entire scientific and engineering research community.’ Remarkable as that omission is when viewed from a systems design perspective, perhaps it is readily understandable as a rhetorical strategy: to focus upon the difficult and all too familiar questions posed by the human organisation components of the system

²² (2003) A-1–2.

²³ (2003) A-3.

undoubtedly takes something away from the construction of a persuasively enthusiastic case for devoting a very substantial amount of funding to its technological elements. The principal problem with this, however, is that to say ‘well, we can always jump off that bridge when we come to it’, and then to hasten onwards, is more often than not a self-fulfilling strategy.

2. THE INSTITUTIONAL CONTEXT OF SCIENTIFIC COLLABORATION

This chapter makes recommendations about how appropriate institutional arrangements, and legal contractual arrangements in particular, might be established for collaborative e-Science. But, in constructing and seeking to implement such arrangements it is important to understand both the informal institutional conditions under which collaborative scientific research projects are organised and conducted, and the specific character of the legal issues that will arise in the organisation of e-Science projects. A few, key aspects of the institutional settings in which arrangements for collaborative e-Science projects are required can have a major effect in determining the success of such undertakings. As these have strong implications for the approach that underlies the recommendations put forward in Part 3, it is important to lay a basis for the latter by reviewing these features of the current ‘institutional environment’.

2.1 Complexities of the Current Institutional Environment

A first and quite important point to notice about academic communities today is that they have been undergoing rapid changes that have left the norms of professional behaviour far from uniform, and in a state of flux. In particular, the scientific communities traditionally had similar (albeit differentiated) norms for the attribution of credit and responsibility for collaborative research. These traditional norms fostered the dissemination of scientific information because the primary incentive for individual researchers (status and recognition within the scientific community), constituted incentive for them to disseminate widely and to accept responsibility for research results.²⁴ A relatively secure career

²⁴ For an account of the classic sociological treatment of the norms of academic science communities, and the modern economic analysis in information-theoretic terms of the

structure also meant that researchers had little incentive to distort or to falsify results because the risk to reputation outweighed potential short-run gains. Yet, the increasing uncertainty of scientific careers has led to more disputes about the attribution of credit and responsibility for research findings. Major journals such as the *Journal of the American Medical Association* and organisations such as the *International Committee of Medical Journal Editors* now have explicit policies regarding the attribution of authorship and responsibility in contexts of collaborative work.

Additionally, universities and other public research organisations (PROs) have responded to government policies urging them to cooperate with business firms sponsoring university-based R&D projects, as well as to make efforts of their own to capture value from academic research results. The resulting pursuit of intellectual property rights, and their exploitation through licensing or the creation of university-owned ‘start-up’ enterprises, has worked to undercut the traditional incentives to rapidly and fully disseminate research findings and methods.²⁵ This weakening of traditional ‘open science’ incentives to claim priority of discovery (rather than securing property rights) has been accompanied in some quarters by the erosion of older normative structures. There is considerable heterogeneity of belief among some communities, most evident among the life sciences, as to whether or not the prime obligation of academic scientists remains co-operation for the advancement of knowledge, or the pursuit of research geared toward

functioning of the resulting behavioural incentives and constraints, see P Dasgupta and P A David, ‘Towards a New Economics of Science’ (1994) 23 (1–2) *Research Policy* 487–521; also, P A David, ‘The Economic Logic of ‘Open Science’ and the Balance between Private Property Rights and the Public Domain in Scientific Data and Information: A Primer’ forthcoming in National Research Council, *The Role of the Public Domain in Scientific Data and Information*, (2003) Washington, DC: National Academy Press (SIEPR Policy paper No 02–030, Stanford University, March 2003 <<http://siepr.stanford.edu/papers/pdf/02–30.html>>).

²⁵ Typically, this has involved a focus on licensing the use of trade-marks and logos, and university-owned patents, copyrights (and more lately database rights), as well as arrangements assigning patents to start-up ventures in which the university takes an equity position. But, more recently British universities are being encouraged by Government funding and policy initiatives to develop a wider array of so-called ‘Third Stream’ activities – i.e., those involving the commercial provision of knowledge products and services (other than teaching and research within their individual institutional purview). Moreover, in some Government circles it is viewed as not only appropriate but imperative that universities have long-term strategies for developing and managing their ‘knowledge assets’ so as be better able to engage in ‘Third Stream’ revenue generating activities. See Jordi Molas-Galant et al, *Measuring Third Stream Activities: A Report to the Russell Group of Universities* (April 2002) SPRU, University of Sussex.

profitable commercial innovations – including those from which they can expect to benefit personally.²⁶

These new trends have not, however, been advancing with uniform strength across all areas of scientific endeavour, institutions or geographical regions. They are, for example, quite evidently far more pronounced in the life sciences, and particularly so among departments engaged in molecular biology and genetics than among departments of theoretical and experimental particle physics. The uncertainty this creates would appear to imply that formal legal rules allocating responsibility for, and the outcomes of, collaborative projects are more important now than ever. Certainly academic researchers are increasingly aware that the law has the power to impinge upon their work, and many are sensitive to the existence of a disjuncture between the norms upon which the law operates and those that have traditionally governed collaborative science.

2.1.1 The Balance between Informal and Formal Governance Mechanisms

Individual scientists, however, have varying degrees of commitment to the traditional norms and in any case lack an obvious forum in which to express dissatisfaction with the law. Moreover, they differ in their interest and talent for handling the administrative aspects of scientific projects. Such differences notwithstanding, most working scientists tend to express impatience with, if not disdain for the effort to formal rules, norms and standards of individual conduct among researchers – even if they acknowledge that this may be necessary to create an appropriate institutional context for the conduct of a new collaborative undertaking.

These attitudes reflect in some part the shared expectation that relationships among scientific peers and co-workers can be governed by the incentive compatibility of co-operative consultative processes (where

²⁶ See, for example, Jason Owen-Smith and Walter W Powell, 'Careers and Contradictions: Faculty Responses to the Transformation of Knowledge and its Uses in the Life Sciences' (2001) 10 *Research in the Sociology of Work* (Special Issue on *The Transformation of Work*, edited by Steven Vallas) 109–40; Rebecca S Eisenberg, 'Bargaining over the transfer of proprietary research tools: Is this market failing or emerging?' in R Dreyfuss, D L Zimmerman and H First (eds) *Expanding the Boundaries of Intellectual Property* (2001); J P Walsh, A Arora and W M Cohen, 'Research Tool Patenting and Licensing and Biomedical Innovation, (December 2002) in *The Operation and Effects of the Patent System*, (Report of the STEP Board of the National Research Council, National Academy of Sciences), forthcoming from National Academy Press in 2003.

the important games are strongly ‘positive sum,’ and there will be a potential for significant damage to individual’s reputation if they are seen to have defected from cooperative play). Impatience with efforts to articulate norms also may stem, in some part, from the supposition that the asymmetry of power relationships within scientific workgroups is well recognised by all the participants; and is understood by them to provide a reasonable enforcement mechanism to resolve the normal run of internal disagreements – for example, by appeal to the authority of the project director. Nevertheless, the introduction of new actors (either in the form of another, collaborating workgroups, or administrative representatives of host institutions) readily can de-stabilise those internal governance mechanisms.

Proposals for new collaborative arrangements, however, introduce the possibility of new incentive structures, which also may be inherently destabilising. Excessive interest in the details of the administrative arrangements might not only be considered a diversion from the core scientific activity, but a threat to the trust upon which scientific collaboration depends. It is important for the success of a collaborative project that the scientists involved understand the broad terms upon which it is to be conducted. At the same time, exposure to a lawyer’s attempt to anticipate all the conceivable potential situations leading to collaborative failure or other unwanted outcomes, and to provide for remedies and mitigating procedures *ex-ante*, may negatively impact the collaborative spirit of the research partners.

Thus, while it surely is salutary to clarify ambiguous social norms, and to reinforce certain professional standards of conduct in situations where these are found to have undergone unwanted erosion, it could be quite counterproductive to attempt to devise and complete detailed contract-like regulations for the internal governance of relations among academic researchers. Such ‘codes’ may have the perverse effect of inducing researchers to think narrowly in terms of legal rights and obligations, and to resort to the often cumbersome machinery of the law in order to resolve disputes among colleagues, or conflicts with university administrators. For example, in the face of unforeseen contingencies, participants in a collaborative project may stand by the letter of the contract rather than co-operate towards averting escalation of the conflict into an organisational crisis. Equally, conscious of having the option of recourse to legal means of protecting their interests, they may

be less concerned to look for ways to remove the source of the problem – say, by rearranging features of the project’s work programme. Explicit preparations to deal with ‘collaboration failure’ by mobilising external (legal) resources, can in this way ‘crowd out’ individuals’ voluntary actions that would render the collective effort more successful. Researchers who feel that no or little trust is placed in them may be more likely to behave in ways that are inimical to the success of the collaboration.²⁷ Evidence from experimental economics and field studies shows that the introduction of explicit contractual incentives can weaken or entirely vitiate the effects whatever intrinsic impulses or social motivations might otherwise be sufficient to elicit cooperative behaviour on the part of the actors.²⁸

The same principle also has a bearing upon the approach to inter-institutional contracting where legal agreements are required. In designing an institutional framework for e-Science collaboratories, a light touch approach may be required to prevent all remnants of the ‘open science’ ethos from being ‘crowded out’ even from the transactions among academic institutions.²⁹ Certainly, when considering the respective roles played by formal contractual agreements and informal norms and understandings regulating the interactions among members of research communities, one should not suppose that these are perfectly complementary, or even strictly ‘additive’. At the margin, each

²⁷ See, for example, M Bacharach, G Guerra and D J Zizzo, *Is Trust Self-Fulfilling? An Experimental Study* (2001) Oxford University Department of Economics Discussion Paper No 76 <<http://www.econ.ox.ac.uk/Research/WP/PDF/paper076.pdf>>, and G Guerra and D J Zizzo, ‘Trust Responsiveness and Beliefs’ (2003) *Journal of Economic Behavior and Organization* - the Discussion paper version available at <<http://www.econ.ox.ac.uk/Research/WP/PDF/paper099.pdf>> (in press, forthcoming in 2003).

²⁸ See, for example, S S Frey and R Jegen, ‘Motivational Interactions: Effects on Behavior’ (2001) 63–64 *Annales d'Economie et de Statistique* 131–53; U Gneezy and A Rustichini, ‘Pay Enough or Don’t Pay at All’ (2000) 115 *Quarterly Journal of Economics* 791–810; and E L Deci, R Koestner and R M Ryan, ‘A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation’ (1999) 125 *Psychological Bulletin* 627–68. For directing US to this pertinent literature, the writers are indebted to Daniel Zizzo, of Christ Church College, Oxford.

²⁹ Further research undoubtedly would be required to assess how light this ‘light touch’ would have to be in various situation, in order to minimise the displacement of informal understandings in contexts where those are more supportive of fruitful on-going inter-institutional co-operation.

may provide a substitute for the other. The problem is that one rarely can know, *a priori*, how deep that margin is, and whether the introduction of a requirement to enter into formal legal contracts may in effect displace, or degrade the effectiveness of informal governance mechanisms. This is an observation that can be formulated in rather general terms, and, indeed, economic analysts frequently made the point that external regulatory provisions mandated by government may ‘crowd out’ the provision of less formal governance arrangements among the agents involved.³⁰

The implications of the foregoing are quite straight-forward. Firstly, the parties seeking to establish a mutually beneficial collaborative research project have both incentives and capabilities to start the process on their own, in a ‘bottom up’ fashion. Beginning without anything like a complete and explicit set of governance arrangements provided by legal contracts (which would carry external ‘third party’, enforcement provisions), they undoubtedly will quickly enter into some informal discussions on key issues: the division of research responsibilities, the arrangements for access to data-streams while the research is in progress, and afterwards; also high on the agenda for discussion will be the project’s publication plans, and the general ‘collective policy’ vis-à-vis intellectual property rights claims to such results as may be anticipated.³¹

³⁰ A very simple illustration of this very general point arises in discussions of the ‘moral hazard’ problem that is alleged to have been created by government programs that insured depositors in mutual savings banks in the US; lulled to a sense of security that they would be protected from losses due to the bank’s inability to pay, the depositor-shareholders paid little attention, and so left unchecked the unsound loans and fraudulent transactions that were made by the executives who managed the affairs of many of those federally chartered institutions. A rather more subtle point to notice is that the introduction of formal prescriptive arrangements governing behaviour – such as contracts – in circumstances where there is uncertainty, and consequently less than complete information, means that such contracts necessary will be incomplete. At best, they can specify *ex ante* features of the process, or procedure that the parties are to follow in coping with unanticipated events affecting their enterprise.

³¹ See, for example, S Hilgartner and S I Brandt-Rauf, ‘Data Access, Ownership and Control: Toward Empirical Studies of Access Practices’ (1994) 15 *Knowledge* 355–72; S Hilgartner, ‘Access to Data and Intellectual Property: Scientific Exchange in Genome Research’ (1997) *Intellectual Property Rights and the Dissemination of Research Tools in Molecular Biology* 28–9 (summary of a Workshop held at the National Academy of Science, Washington DC, 15–16 February 1996); S Hilgartner, ‘Data Access Policy in Genome Research,’ in A Thackray (ed), *Private Science* (1998), on the arrangements for access and control of data-streams that emerged among the teams participating in the Human Genome Project. These internal data-sharing agreements

Secondly, there is reason to expect that there will be latent or manifest grounds for members of such groups to devote time and effort to activities that are likely to reinforce cooperative attitudes and behaviours among the participants. Those efforts include an array of 'natural' social contacts (dyadic transactions as well as collective assemblies) which facilitate monitoring of the personal dispositions and social attitudes of colleagues, and can contribute to raising levels of trust and trustworthiness. To engage in non-committal speculative discussions that explore the possibilities of successive projects of potential mutual benefit, the formation of which would be contingent on the successful outcome of the immediate prospective collaboration, would have a similar function. They serve to embed what otherwise might be construed as a 'one-time transaction' in a 'super-game' that features repeated play. The pay-off structure of the latter form of game tends to induce (rational) participants to defer defections from co-operation, even if it remains unwarranted to assume that acts of self-interested opportunism at the expense of the rest of the group will therefore have been foresworn by every one of the players.

Therefore, it is not unreasonable for scientists to be disposed to avoid, to the utmost extent possible, both efforts to codify administrative rules for research management, and the framing of legal contracts for governance of collaborative projects. In addition to the analytical considerations already reviewed, there is empirical experience in the field of contract law that even in business affairs many parties are reluctant to use the law in the planning of their relationships, because they fear that it will harm the collaborative nature of purely commercial relationships.

The upshot is that the desire of scientists not to become embroiled in such administrative and legal matters ought to be respected in determining appropriate governance structures for collaborative projects, even when these require the provision of legal contracts. From this position it follows that what scientists would find most helpful in pursuing research by means of multi-party collaborations is a menu from which to select 'ready-made' solutions to the more commonly occurring specific problems in such enterprises; and the option to reconfigure the

may be contrasted with the observations of the weakening in recent years of the ethos of general data-sharing in the biomedical sciences (noted above, in section 2.1).

elements of those solutions to fit the idiosyncratic requirements of their particular circumstances.

The indicated solutions can be of two sorts. For intra-project relationships among employees of a given administrative unit (which includes an entire hierarchically administered institution such as a research institute or a university), the menu should emphasise reliance upon informal, peer-enforced norms of conduct, and alternative procedures for dispute resolution. But, the governance arrangements pertaining to research relationships that involve collaborations across institutional boundaries, whether with other PROs or with business firms, will require legal contracts; the menu of alternative contractual clauses in that case must feature the array of provisions from which a comprehensive agreement can be constructed.

2.1.2 Conflicting Interests in Institutional Collaborations and Partnerships

The principal *legal* actors in the establishment of a collaborative research project generally are not the researchers involved, but, as has been noted, the institutions by which they are employed. This is both appropriate and presents real dangers. On the one hand, it is appropriate that scientists should be relieved of the burden of establishing and maintaining the infrastructures of collaborative work. On the other hand, there is a danger that the farther the creation and maintenance of such structures is placed from the activities of collaborative science itself, the greater is the likelihood that those structures will reflect the interests of actors other than the collaborating scientists, and worse, actors whose goals may be inimical to the effective conduct of the project.

Universities do have multiple and conflicting institutional interests in relation to collaborative scientific research. They may find it in their interest to foster collaborative science as a means of supplementing their intellectual capabilities, enhancing the institutional research reputation of their schools and departments, and attracting ‘star’ researchers and research funding – both in the near term and in the longer run. In some areas, such as the European Union Framework Programmes, contractual participation in multi-institutional networks is the *sine qua non* for obtaining external funding, and, increasingly in recent years public

research agencies and private foundations have encouraged the organising of collaborative projects of that kind.³²

At the same time, however, universities also have acquired a stake in the promotion of some government policies that may well turn out to impede collaborative research in science and technology. Universities in the United Kingdom, for example, have adopted increasingly comprehensive policies of asserting ownership of any intellectual property in material produced by their research and teaching staffs. In some instances it is true that clear policies regarding the ownership of intellectual property may facilitate the resolution of disputes between collaborators. In other circumstances, however, the potential conflicts of the IPR claims asserted by the different institutions that employ the would-be collaborators (i.e., the research scientists) can create impediments to the formation of scientifically promising projects. Indeed, the prospective transactions costs incurred in trying to resolve what are in essence distributive conflicts among the prospective participants may even frustrate formation of research undertakings that are likely to yield high rates of return to the coalition as a whole.

Universities are particularly complex organisations, however, and their key administrative leaders typically are well aware of the multiplicity of distinct missions that society expects them to pursue concurrently, so that the balance of priorities among those missions often is less than clearly defined within each institution.³³ In some respects this degree of

³² For discussion and analysis of endogenous coalition formation in response to such programs, see P A David and L C Keely, 'The Economics of Scientific Research Coalitions: Collaborative Network Formation in the Presence of Multiple Funding Agencies' in A Geuna, A Salter and W E Steinmueller (eds), *Science and Innovation: Rethinking the Rationales for Funding and Governance* (2003).

³³ In Britain, the Russell Group of Universities has recognised this in a recent expression of concern that Government core funding for universities might be associated with easy-to-measure features of entrepreneurship, technology transfer, commercial knowledge service provision, and still other so-called 'Third Stream' activities – at the expense of the many other forms of interaction between universities and the economy. 'This concern is particularly acute among leading traditional universities that value the close integration of teaching and research, that operate across all the disciplines, and that engage with society in very many diverse ways that include, but are not limited to economic transactions. Some are undertaken to achieve a directly financial outcome, while most are promoted for their wider, out-reach and often long-term benefits. Furthermore, there is a broad spread of missions within the British University sector.' See: <http://www.clo.cam.ac.uk/3rd_arm_metrics.htm>. It is a sign of the times that the statement of the Russell Group, rather than focusing concern upon the possible sacrifice of the universities' creditability and capacity in a wider set of non-commercial social interactions,

ambiguity is a source of flexibility and accommodation to special circumstances, but it also may permit formal contractual arrangements for co-operative research among them to becoming ‘snagged’ on points of importance to one or another of the participants – even when those matters have little to do with distributive conflicts, and consequently are difficult to resolve by arranging inter-institutional ‘side-payments’.

Quite understandably, legal departments operating within universities must maintain a strong professional commitment to protecting the institution by limiting its exposure to the particular risks of collaboration, as well to the losses that may ensue from the conduct of research by their own employees and contractors. Sometimes, therefore, foreseeable and uninsurable risks the individual institutions would have to bear will appear too large in comparison to the uncertain benefits they might derive, and so it will turn out to be far easier to negotiate ‘safer’ collaborative projects, or ones with more readily calculable future income-streams. Furthermore, in response to the thrust of recent government policies promoting the search by universities for income from the commercialisation of their so-called ‘knowledge assets’, the efforts of technology transfer offices and intellectual property management organisations are directed towards fixing the terms of collaborative research projects so as to augment the flow of income to their respective institutions. Of course, this pits each against the similar interests of the other collaborating institutions, and likewise against the business companies whose participation (and sponsorship) is predicated on obtaining a satisfactorily large share of the prospective economic returns.

One should not underestimate the seriousness of the difficulties that have thus been created for university officers who are given responsibilities for negotiating true inter-university agreements that have to resolve conflicts over the division of prospective proceeds from the commercial exploitation of research findings. The same conditions also may give rise to tensions between groups of academic researchers who are keen to participate in a particular collaborative project and their

focuses on the need for methodologies to allocate government support for the diverse array of ‘Third Stream’ activities in which they typically engage, and therefore to protect their institutions position in the face of potential competition from more specialised claimants for ‘Third Stream’ support.

respective university legal counsels, who are equally intent upon limiting (as far as is possible) their own institution's exposure to liability and other legal risks, as well as asserting its claim to the largest possible flow of material benefits that the project might yield. The immediately perceived interests of the university as a legal corporate entity may involve asserting provisions that do not necessarily advance the interests of the researchers, and indeed would be enforced at their expense. A more subtle effect that, as has been noticed, can have a corrosive effect on the trust necessary for successful research collaboration, is the well-intentioned effort of a diligent lawyer to render explicit *ex ante* all the things that conceivably could go wrong, and all the forms that betrayal of trust that the collaboration might sustain.

Decision-making becomes particularly complex when the interests of the parties diverge (let alone conflict) and no one set of specialised actors understands all the issues and is in a position to balance the potential risks of going forward with the collaboration as it has been designed by the researchers, against the risks of losing the prospective benefits by imposing a different set of arrangements that would frustrate the research itself, or weaken the incentives of the participants to behave co-operatively. Numerous examples might be cited in this regard, but a familiar comparatively benign illustration of the general problem may be seen in the situation of the specialised service offices that most research universities have found it necessary to establish in order to facilitate the transactions with external sponsors of research.

Typically, the 'sponsored research office' in a US academic institution (and its counterpart, the 'university research services' office in Britain), develops familiarity and expertise in regard to the panoply of regulations and requirements that funding bodies impose upon applicants, and recipients of awards; and equally specialised skills in anticipating the issues that will arise in negotiations with corporate research partners. The performance of these organisations is likely to be gauged primarily in terms of the volume of funding that their respective institutions receive from public (including charitable foundation) grants and contracts, and also from research partnership agreements concluded with business firms – not the satisfaction of the researchers, or the scientific and scholarly productivity of their sponsored research projects. University research services officers often have qualifications and practical professional experience in the law, but their role is quite distinct

and in a sense more demanding than that of university legal counsel *per se*. They are thus able to provide intermediation services that greatly reduce the burden of ‘negotiation and administration’ upon the researchers. But, with expertise and the quest for efficiency comes a tendency on the part of such offices to promote compliance with standardised contractual formulae, to avoid undertaking to provide novel or highly ‘customised solutions’ that may better fit the needs of a particular research project. The main institutional problem posed by such solutions is that they are unfamiliar, and hence all too likely to occasion time - and attention - consuming special negotiations with other departments within their own university, as well as with the external agencies.³⁴

As a specialised intermediary the sponsored projects office is thus predisposed by its own incentive structure to implement, rather than to question the need for contractual provisions that are asked for by diligent university solicitors – functioning in their specialised domain to protect the institution from the entire array of harms to which it may be exposed by a proposed research agreement. Thus, it is the extraordinary ‘university research services office’ that can be expected to take upon itself the role of serving university researchers as an ‘agent’ serves a ‘principal,’ and a professional firm of lawyers serves its clients. At the same time, although it occupies the intermediary position upon which converge the varieties of diverse and possibly divergent interests within the university in regard to particular proposals for collaborative research, these service organisations have not been given explicit discretionary authority to strike a balance among those contenting interests.

For the individual researchers who have initiated a collaborative proposal, matters are made more complicated and potentially more difficult by the fact that the principal contracting parties, legally speaking, will not be their other scientific colleagues but, instead, those colleagues’ respective institutional employers. This really is a two-sided problem, because the success of most collaborative projects will depend upon the work of individual researchers and research administrators

³⁴ In the latter context, it is quite possible that the presentation of unusual (anomalous) contract proposals may be read by their counterparts outside the university as a symptomatic of deficient professional competence, or lack of authority within the institution they serve; whereas, just the opposite interpretation would be more forthcoming were they perceived to be acting on behalf of a legal client – in this case the researchers, or their university.

with scientific expertise. The peculiar employment structure of academic institutions, with their tradition of academic freedom, and the increasing mobility of scientific researchers, means that the contracting university often has little control over its ability to deliver the services that it promises in the contract to establish a research project. It may well be that if the incentives of universities to enter collaborations are further increased in the future, there will be corresponding efforts to alter the contractual relations with their employees so as to more closely emulate business corporations – particularly in respect to the power of the latter to direct and control the participation of employed researchers in designated co-operative projects involving other organisational entities.

While there will be obvious practical difficulties in pursuing such a restructuring of university employment relations, it is not evident today that the idea of would be rejected quickly, either as inappropriate or as difficult to reconcile with the ethos of scholarly autonomy. Increasingly, universities in the United Kingdom and elsewhere are coming to regard their members (faculty and student members alike) as ‘knowledge assets’. The very term suggests that those responsible for the disposition of university assets should be able to exercise more complete and exclusive control over faculty members’ activities, so as to better deploy them in pursuit of greater revenues from the intellectual property and commercially valuable ‘knowledge services’ they are able to generate. In most cases there is little likelihood that the incremental revenue streams thereby captured from the intellectual property arising through the assigned work of employees and students will materially alter the institution’s financial situation.³⁵ Yet, as economists are quick to point out, resource allocation decisions are determined ‘at the margin’. Consequently, it is quite frequently observed that the prospect of the university achieving a comparatively small financial benefit will elicit the expenditure of significant administrative effort in altering long-established policies and operating rules. Likewise, the prospect of the university being exposed to a low-probability and low-cost risk, or of

³⁵ See David R Mowery et al, ‘The Growth of Patenting and Licensing by U. S. Universities: An Assessment of the Effects of the Bayh-Dole Act of 1980’ (2001) 30 *Research Policy*; Richard Nelson ‘Research and Technological Progress in Industry-An analysis of the American Experience’ (International Symposium on Economic Development through Commercialization of Science and Technology, Hong Kong 2002).

having relinquished some small gain to a ‘partner’ institution in a particular project, will all too likely result in efforts by the institution’s solicitors to forestall such outcomes by imposing blanket *ex ante* restraints upon the actions in which individual employees participating in a collaborative project are allowed to engage.

The three dynamics that have been reviewed in this section are at present only partially understood. This is a problem, because the viability of various possible institutional arrangements to support e-Science ventures depends to some extent upon knowledge of the situation in which many collaborators in a designated research area are likely to find themselves. More information, therefore is needed not only about the direction of future institutional changes affecting PROs, but about current individual experiences with informal arrangements, and with the ways that formal legal rules are presently being utilised by co-operating institutions. As information of this sort would be gathered more or less automatically as a part of the process recommended in the chapter’s fourth and concluding part, further discussion need not be pursued at this point.

2.2 The Institutional Design Challenge: Analytics

Seen in properly broad perspective, the global e-Science system design challenge is one of finding the set of technical *and* social and mechanisms that will provide the collaboration facilities, incentives and controls needed for human-machine research organisations to emerge and function efficiently in cyberspace. Inevitably, this challenge will present itself in many different and very specific contexts in the course of the co-evolving interdependent adaptation of the system’s institutional and technological components.

It follows that close attention should be devoted to the requirements of both sets of components, and to exploring the possibilities of applying ingenuity and resources to achieve innovative solutions in the institutional as well as the technological domains. This will be a critical strategy not only for the long-term success of e-Science programmes, but for securing the potential benefits of the contribution they can make to enhance the effectiveness of collaborative work in many other areas of human endeavour.

2.2.1 *'Collaboration Games': Players, Interests and Rules*

To make headway towards that worthy objective, it is important to try to more clearly delineate the fundamental sources of the organisational and institutional problems that need to be addressed by PROs when defining the terms, rights, responsibilities and powers of the legal parties to research collaborations. 'Collaborating parties' is of course a blanket term that covers at least four categories of entities/actors. For the purposes of analysing the sources of such conflicts, the following should be distinguished:

- the research scientists as individuals;
- research units that have formal governance arrangements (departments, laboratories, and other consortia bound by agreements among the participating researchers);
- host institutions with whom the researchers are connected by contracts, and through which they may receive financial or other material support (universities, public institutes, foundations and trusts, private partnerships and corporations);
- public and charitable funding bodies, and private business organisations that furnish material support to PROs for the conduct of research (and related training) activities.

These are presided over, of course, by national and international entities that may exercise primary or derived regulatory jurisdiction over both the individual researchers qua citizens (in the case of national governments) and their host institutions.

For the purposes of a general discussion, and also in some specific contexts, it is helpful to further simplify matters by consolidating the foregoing list into two categories, the first pair of parties being lumped under the heading of 'research collaborators' and the second aggregated into 'institutional partners/hosts'. Seen from that highly stylised perspective, there are two sets of core difficulties of designing supporting institutional arrangements for collaborative projects. These may then be succinctly characterised as arising in the first place from the

imperfect alignment of the interests of the research collaborators, on the one hand, and the institutional partners, on the other hand; and in the second place, from the existence of mal-alignments or outright clashes of interests among the institutional partners. In some significant degree, the second class of difficulties also may contribute to the tension between the researchers whose goals impel them towards collaboration, and their respective institutional hosts for whom the terms of such coalitions may be problematic. Only in what economists term ‘perfect teams’ is it appropriate to ignore the consequences of conflicting goals and interests among the parties, but in the world of human agents the ‘perfect team’ is a conceptual device, not a reality.

Incentives for competition and collaboration are important for both types of actors in e-Science collaborations. But, it must be recognised that the point where those two forces would be ‘naturally’ balanced is not the same in the typical case of collaborating researchers as it for the institutional entities (i.e. public and private corporations and institutes and universities) that enter into a research partnership or consortium. The situation within scientific work-groups is usefully distinguished from that which typically holds among participating institutions, and can be examined first.

Within scientific work-groups situated in academic milieux, it generally holds true that the ‘open science’ ethos and traditions of scientific co-operation among researchers forms at least the point of departure (or default position) for the ‘bottom up’ organisation of collaborative activities. This remains the predominant orientation, even though norms of co-operative behaviour are strained by rivalries for scientific recognition and reward; and it continues to be useful as a first-order idealisation of complex situations, the existence of considerable variations among the local *mores* characterising ‘open science’ communities in different fields of inquiry notwithstanding.

What the future holds in this regard remains unclear. A more elaborate and nuanced account of the current situation in academic science would emphasise the respects in which norms of professional behaviour and institutional policies are in flux. It is uncertain that the traditionally prevalent disposition in favour of scientific co-operation will be able to withstand the pressures from the newly ascendant spirit of ‘academic entrepreneurship’. Similarly, at the institutional level it is far from

obvious that commitment to the collective goal of ‘the pursuit of knowledge’ will continue to guide the policies embraced by a majority of leading research universities. In many places it already has ceased to prevail in the face of the instrumental emphasis placed by public funding agencies upon the wealth-creation function of knowledge, and the growing legitimisation of the pursuit by researchers of personal wealth through ownership of intellectual property.³⁶ Quite obviously, these are important issues not only for the scientific communities involved, but also for society at large. Moreover, they are issues whose ultimate outcome can be powerfully shaped by the effects of myriad, seemingly small decisions about the technological and institutional infrastructures of e-Science.

Among the institutional partners of a collaboration, by contrast with the scientists carrying out the research (and for whom publicly funded research universities may be said to serve as hosts), the predominant natural orientation lies more towards competition than co-operation. This generalisation may seem paradoxical, but at a fundamental level it follows from a simple contrast in motivations. The primary objective of collaborating with other researchers is to gain access to the immediate data and informational fruits of each other’s knowledge. But, data and information are public goods that may be exploited by all the collaborators in their respective research endeavours – without becoming exhausted or in any respect diminished. By contrast, the driving motivation for the corporate entities to enter a research partnership or coalition (*qua* university or *qua* business corporation) is to gain access to material benefits that do not possess ‘public goods’ properties, so that the existence of opposing interests among them over the division of the pie is ineluctable. When university/hosts are moved to become partners in a multi-institutional project, each is likely in some degree to be responding to the influence of derived motives for co-operation – whether it is to accommodate the scientific work of academic collaborators, or thereby to gain the overhead funding, or the possible payoffs in prestige and command over material resources that a

³⁶ See, for example, Jason Owen-Smith and Walter W Powell, ‘Careers and Contradictions: Faculty Responses to the Transformation of Knowledge and its Uses in the Life Sciences’ (2001) 10 *Research in the Sociology of Work* (Special Issue on *The Transformation of Work*, edited by Steven Vallas) 109–40.

successful project might bring. But such co-operative and accommodating motivations tend to be tempered, and sometimes overridden by the attention that the institutions accord to protecting and promoting their respective individual, and essentially competing interests.

Where universities are involved as the proximate corporate parties, the primary impetus towards facilitating co-operation derives largely from the interests of the researchers themselves. This impetus also is likely to be reinforced by the terms on which funding may be obtained from the public sponsors of those research projects. In addition to the prospective division of whatever 'pie' will thus be made available to the coalition, the negotiated terms of each institution of higher education's (HEI) immediate relationships with its institutional partners also must be shaped by its entirely understandable concern to manage the uncertainties surrounding the conflict-laden aspects of these partnerships. Moreover, in the United Kingdom and other countries, the policies of governments now encourage higher education institutions and other non-profit research organisations (such as research hospitals) to try to exploit – whether for themselves or for their own private sector partners – the intellectual property arising from the contractually specified activities engaged in by their employees, and even their students. It is especially relevant for e-Science (as it is also for e-Learning) that these legal and administrative arrangements for institutional appropriation of the benefits of new knowledge have been focused particularly upon works created in the form of digital information.

Furthermore, the United Kingdom's government (along with other states in the European Union) appears to be pursuing regional development strategies based upon the formation of business 'clusters' in the environs of publicly supported research institutions, and the promotion of industrial enterprises founded by the licensing of intellectual property generated within the university, or in university-industry research partnerships. University leaders are understandably responsive to the prospects of direct and indirect benefits they may derive by successfully fulfilling this development role, and are thereby induced to aggressively seek to expand and exploit their intellectual property portfolios. It might be noticed, however, that presently there is

very little in the incentive structure that enjoins United Kingdom institutions to consider the burdens that their policies in the latter regard can impose upon the work of their own research and teaching activities; and, *a fortiori*, upon the parallel activities carried on by colleagues at other institutions in the UK. Indeed, nothing in the present incentive structure requires a non-profit, publicly subsidised HEI to consider whether its intellectual property licensing strategies, and its promotion of university ‘start-ups’, will be likely to impose burdens upon the innovation activities of private commercial firms.

2.2.2 Cooperation v Competition When Assets are Complements: Some Guidance from Economic Analysis

For the purposes of this chapter, the foregoing conditions may be taken to characterise the prevailing and prospective state of affairs. They thus describe the pertinent environment within which practical institutional arrangements for the conduct of collaborative e-Science will need to be achieved. It may be argued that societal interests would be better served by promoting more active co-operation among the public sector entities, whereas the recent thrust of government policy (in the United Kingdom particularly) has been to encourage higher education institutions to form co-operative relations with business companies while competing ever more vigorously with each another for command over material resources. But, in the likely absence of a radical policy reversal, the task of creating an appropriate institutional infrastructure for e-Science must be one of devising mechanisms that are better able to strike the socially most efficient attainable balance between the proximate goals of collaborating researchers and the immediate objectives of their host institutions.

The rationale for choosing this particular ‘second best strategy’ is not that scientific collaborators should be accommodated because they are intrinsically good, more selfless, and more worthy of trust than other members of society. Quite the contrary. Precisely that reality must be taken into account by the internal governance structures of the institutions that employ human agents and support their activities, whether scientific or other. The stance adopted in here derives from a different consideration altogether: namely, a recognition of the larger beneficial consequences for society that derive from the strong imperatives for co-operative behaviours in the field of scientific inquiry

(and equally in other complex forms of cultural production). This rationale follows from some fundamental propositions in the economics of knowledge and information, and their application to the analysis of the role of implicit and explicit collaborations scientific and technological research.

Modern economic analysis of the production and distribution of reliable knowledge proceeds from the widely accepted proposition that the advancement of scientific knowledge and the technological progress are intertwined cumulative processes. Both are synergised by complementarities among the data and informational inputs that enter into systematic research activities and which, in turn, are the primary outputs of those activities.³⁷ Co-operative sharing of knowledge resources is well known to be the most efficient allocational scheme for the production of goods when different agents hold complementary inputs. This holds for normal economic goods, and is true for data and information – true, *a fortiori*, as these possess special properties that render them akin to ‘pure public goods’. The latter typically are integral and hence indivisible; yet they are infinitely expansible, being useable repeatedly and concurrently by many agents without becoming depleted or otherwise exhausted. Furthermore, to utilise such goods while denying others access to them is typically costly to arrange, even when it is technically feasible.

Formal analysis based upon the foregoing propositions has established an important result that also (satisfyingly) receives confirmation in behavioural experiments.³⁸ Where the complementary elements required

³⁷ For a more extensive presentation see, for example, P A David, ‘The Economic Logic of ‘Open Science’ and the Balance between Private Property Rights and the Public Domain in Scientific Data and Information: A Primer’, forthcoming in National Research Council, *The Role of the Public Domain in Scientific and Technical Data and Information: A Symposium* (2003) National Academy Press <<http://siepr.stanford.edu/papers/pdf/02–30.html>>.

³⁸ See James Buchanan and Yong J Yoon, ‘Symmetric Tragedies: Commons and Anticommons’ (2000) April 43(1) *Journal of Law and Economics* for a theoretical analysis that makes use of Cournot’s theory of oligopoly behaviour in markets for complementary products. On the latter, see Carl Shapiro, ‘Theories of Oligopoly Behavior’ in R Schmalensee and R Willig (eds), *Handbook of Industrial Organization* (1989) 330–414. Carl Shapiro ‘Navigating the Patent Thicket: Cross Licenses, Patent Pools, and Standard-Setting’ in A Jaffee, J Lerner and S Stern (eds) *Innovation Policy and the Economy* (2003) vol 1, develops the argument that where intellectual property rights in complements are distributed among many agents, compulsory ‘pooling’ of rights and cartel pricing yields a socially more efficient allocation than absolute individual monopoly rights of the sort granted to patent-holders. For experimental validations of

to obtain a valued outcome are controlled by multiple agents, each of whom has the power to exclude others from use of at least one input, the ‘prices’ that these private goal-seeking agents independently will place upon the resources under their respective control will be too high. That is to say, they will fail to consider the upward-cascading effects of the charges they set for access to their respective resource-holdings, with the result that the level of level of production in the system as a whole is sub-optimally low.

Moreover, if there are as well strong complementarities in the use of the outputs for final purposes (such as commercial innovations based upon scientific discoveries or research tools), the dispersal of exclusive ownership (and hence access control) among many parties tends to yield a less-than-socially optimal degree of utilisation of the available information-inputs, and a correspondingly sub-optimal level of consumption of the final goods. Overall, the anticipated results of the envisaged ‘anti-commons’ equilibrium are less socially efficient than those that would be obtained were the knowledge production activity to be organised under the terms of an intellectual property licensing ‘pool’ that pursued a monopoly pricing policy.

The latter is a very strong finding indeed. It is widely accepted that regimes characterised by competitive rivalry generate incentives for cost minimisation and efforts to satisfy the needs of final consumers, whereas monopoly imposes significant inefficiencies, particularly where the product that is monopolised possesses the properties of a pure public good’ – which is the case for information-goods. Hence, finding principles for formally organising e-Science collaborations, and doing so under contractual terms that manage to avoid outcomes that will be ‘worse than monopoly’ is a challenge well worth trying to meet.

propositions about the symmetry and commons and anti-commons forms of market failures, see Charles F Mason and Owen R Phillips, ‘Mitigating the Tragedy of the Commons through Cooperation: An Experimental Evaluation’ (1997) 34 *Journal of Environmental Economics and Management* 148–72; Steven Steward and David J Bjornstad, ‘An Experimental Investigation of Predictions and Symmetries in the Tragedies of the Commons and Anticommons’ *Joint Institute for Energy & Environment Report*, JIEE 2002–07 (August 2002).

2.3 The Formal Legal Context for Collaborative e-Science

It is notoriously difficult to describe exhaustively the legal issues that could arise from the many different types of collaborative e-Science project. Nevertheless, common patterns of legal problem are likely to arise. These concern the legal relationship between the parties to an e-Science collaboration, the material that the parties bring to an e-Science collaboration, the material to which such a collaboration gives rise and the liability of the parties for harms arising from the project. Each of these issues merits separate consideration. That consideration must remain fairly general as the identification of legal issues is very fact-sensitive. But, even general consideration of the likely issues suffices to reveal the complexity of the legal context in which collaborative e-Science will operate. In light of that complexity, it is clear that these are issues that cannot be navigated by individual scientists, and that institutional arrangements similar to those outlined in the final section of this chapter will be essential.

2.3.1 Relationships among the Collaborating Parties

The legal rules that govern disputes between the parties to an on-line collaboration are determined by the nature of the legal relationship between those parties. Three possibilities suggest themselves. First, the parties may be in no particular legal relationship and the general law will determine issues such as the allocation of the fruits of their joint activities. This situation is the most unlikely in all but the smallest and most informal of collaborations. Second, the parties may be in a contractual relationship. If they are, then the terms of that contract will in most circumstances determine the conduct of their relationship and the allocation of its outcomes. Third, the parties to collaboration may establish some type of institutional vehicle for their collaboration such as a joint venture company. This is the most formal way of establishing a collaborative relationship. It has the advantage that it can facilitate the structuring of complex relationships. It also means that individuals and institutions can avoid liability in situations in which a collaborative project may give rise to harm to third parties. However, it entails the maintenance of a system of legal formalities that is likely to make it unattractive to all but the very largest of collaborations.

Of these three possibilities, the most likely one is that the relationship between the parties will be governed by either an express or implied

contract. The question of whether the parties are in a contractual relationship is not simply that of whether they have entered into a written agreement. Even in situations in which the parties cannot point to any such written agreement, the courts may well determine that they are bound by contract either because of an express verbal agreement between them or because the court is prepared to imply a contractual relationship. The freedom that parties have to enter into contractual relations can give them considerable control over the conduct of their relationship and the allocation of its outcomes. But, difficulties may arise; both when collaborators find themselves party to contracts the terms of which they did not expressly agree and, alternatively, when apparent agreements between collaborators are ineffective as contracts.

A first difficulty that arises in relation to the contractual organisation of the relationship between collaborators relates to the issue of when, and on what terms, their relationship is formed. This problem will not arise in circumstances in which the collaboration is based on a written signed agreement which is remade by all the parties to the collaboration each time its terms are amended or the identity of the parties changes. Many collaborations, however, will be organised on a more informal basis. A part of the attraction of e-Science is that it can involve many different parties with different contributions to make who become involved at different stages in a project. Two particular problems are likely to emerge. First, the parties to a collaborative project will often amend their agreement informally during its life. This may cause a problem if one of the parties is promising to do more than she undertook under the original contract, while the other is not. Under English law, the problem of giving effect to such amendments is a textbook difficulty in the law of contract. Other legal systems have less difficulty with such situations and the English courts have been working to find a solution to the problem. Nevertheless, it is one that may well arise in on-going scientific collaborations.

Second, latecomers to a project may be included in its work without entering into an express agreement with all the existing parties.³⁹ In

³⁹ Academic lawyers puzzle over cases such as *Clarke v Dunraven* (1895) P 248 and how rules of contract law designed to reflect the paradigm of two or more parties in an express agreement on terms that govern their relationship throughout its life can be adapted to the context in which parties attempt to join a single agreement at different points in time.

relation to the rules of unincorporated members' clubs, the law has developed methods for dealing with the problem of how contracts governing on-going relationships can be extended to include newcomers. In such contexts it is generally assumed that the members of a club are in contractual relations with one another; that the terms of that contractual relationship are set out in club rules, including procedures for the admission of new members; and that those rules can only be altered either in the manner prescribed in the rules themselves or by the agreement of every member of the club. There is no reason why such an approach might not also be adopted as a means of organising the entrance of participants who joined the collaboration at various points after its inception. This may not be possible, however, if the newcomer were to enter the relationship on a basis that has not been anticipated at the outset of the project. Moreover, even in situations in which the agreement between the collaborators allowed for the admission of newcomers, the terms of the contract among the original members of a collaboration would need to be made clear to the late-comer. The difficulty with contractual analyses of these types is that the outcome to a dispute between the parties, particularly one involving parties to a collaborative relationship who have joined at different points in time, may not reflect the expectations of all the parties. Indeed, such expectations are unlikely to coincide.

A second difficulty will concern the appropriate parties to any written contract that forms the basis of the collaboration. For example, a written agreement may be arrived at among institutions rather than individuals, but a collaboration will require agreements (whether formal or informal) among particular individuals who possess with particular expertise. Difficulties may arise where an institution enters into an arrangement for a particular collaboration, the individual who is central to the collaboration moves institutions and the original institution seeks to replace them and to continue the project with someone with whom the other collaborators are not content. A similar difficulty might arise in the context in which the parties to collaboration assume that they are entitled to the fruits of that collaboration but, in fact, the general law grants first ownership of those fruits to another party. An example of this might be the situation in which individual collaborators assume that they are able to make arrangements concerning the fruits of their collaboration, but at least one party's share falls under the general law to

the institution by which she is employed. There are various ways in which such problems can be overcome, but it is essential that the issue of the most appropriate contracting parties be addressed before the project begins.

This highlights a third potential difficulty with the contractual organisation of a collaborative research project. Where the fruits of a collaborative arrangement constitute the subject matter of one of the statutory intellectual property regimes, the statutory code itself allocates its first ownership. As outlined below, the parties to collaboration are free to arrange between themselves that the normal rules for the allocation of property in such circumstances should be altered. Yet, the process of doing so involves some possible pitfalls. In particular, informal agreements may be difficult to enforce. Take, for example, the situation in which one of the parties to collaboration writes the software upon which the project is based and copyright law allocates first ownership of that software to her. The author of the software may, by agreement, assign her interest in the software to her collaborators or license them to use it. She may even, as a part of the original collaboration agreement, have assigned her interest in any prospective works, including the software that she might write as a part of the project. In the United Kingdom, this is provided for by section 91 of the *Copyright, Designs and Patents Act 1988*. However, to do so she must comply with the legal requirements of such an assignment and under section 90 of the *Copyright, Designs and Patents Act 1988*, these include a requirement that assignments be in writing signed by the assignor. Any agreement that may be entered into as to the allocation of the fruits of collaborative work must not only be clear and made between the appropriate parties, but also meet the formalities required by any applicable law.

A fourth set of difficulties relates to the often-unequal bargaining power of the parties to collaboration. The law is able only in the most extreme circumstances to correct the effects of unequal bargaining power. This is one reason why it is vital that common understandings of the appropriate response to particular issues that arise in planning scientific collaborations should be developed. Scientists may come under considerable pressure from their commercial partners to agree to terms that do not further the goals of collaborative research generally, and to which they would prefer not to agree. To illustrate this point, one may

consider the way that the effects of such asymmetries shaped the outcome of an unusual bargaining situation involving intellectual property, a recent case where no revenues from IP licensing were at stake. The case in question involved negotiations between the publicly funded Globus Project at the Argonne National Laboratories (Chicago) over the form of the ‘open source’ license for the Globus Toolkit software under which major business firms – including IBM and Oracle to be specific – would undertake the distribution of that software package.⁴⁰

Compared with those companies, the Globus Project had very limited capabilities to engage on its own in a widespread ‘free and open source’ distribution of the Globus Toolkit software. It might have done just that under the familiar form of GNU General Public Licence (GNU GPL).⁴¹ But IBM and Oracle were unwilling to expose themselves to ranges of risks they perceived would ensue were they to distribute code under the ‘copyleft’ type of license familiar in ‘free and open source software’ (F/OSS) products released under GNU GPL. Recognition of its limited leverage, and a desire to rapidly establish the GT protocols as

⁴⁰ See David and Spence (2003/2004), Appendix 4 (The Globus Project’s Approach to Software Licensing) for further details and references.

⁴¹ The question of whether or not a computer program qualifies as ‘free and open source software’ (F/OSS) can be approached as a matter of legal definition. (See David McGowan, ‘Legal Implications of Open Source Software’ (2001) 241 *University of Illinois Law Review*.) Copyright law gives developers who write programs the exclusive right to reproduce the code, distribute it, and make works derived from their original work. Copyright holders can grant other parties permission to do such things through licenses, and ‘free’ software or ‘open source’ software refers to software distributed under licenses with particular sorts of terms. The ‘Open Source Definition’ (maintained by the Open Source Initiative, see *The Open Source Definition* version 1.9 <www.opensource.org/osd.html>) provides a convenient and widely accepted reference guide to such licenses. It sets out several conditions a license must satisfy if code subject to the license is to qualify as ‘open source software’. Several well-known licenses satisfy the Open Source Definition, the most widely used and still more widely discussed among them being the GNU General Public License (GPL). Programs distributed under a F/OSS license ‘must include source code, and must allow distribution in source code as well as compiled form; it also must allow modifications and derived (GPL’d) works, and ‘permit them to be distributed under the same terms as the license of the original software’. Such a license does not restrict any party ‘from selling or giving away the software as a component of an aggregate software distribution containing programs from several different sources,’ but the licensee cannot ‘require a royalty or other fee for such sale.’ Much of the attention given to F/OSS development focuses on the GPL’s requirement that authors who copy and distribute programs based on GPL’d code (derivative works) must distribute those programs under the GPL. This requirement is specified in Section 2(b) of the GPL and is referred to as the ‘copyleft’ term.

a Grid standard, appears to have been sufficient grounds for the Globus Project to accede to its industrial partners' demands by working with the companies' attorneys to devise a customised software license.⁴²

The effects of asymmetry of bargaining power are further revealed by considering a parallel case of negotiation, in which a different and less restrictive form of software license was adopted. Sun Microcomputer's Project JXTA has created a set of open, generalised peer-to-peer protocols that allow any connected device (cell phone, to PDA, PC to server) on the network to communicate and collaborate. The source code for Project JXTA has been released under a variant of the Apache software License: *The Sun Project JXTA Software License* is functionally equivalent to the *Apache Software License (Version 1.1)*, with minor changes to reflect the Project JXTA name and Sun Microsystems as the original contributor. In addition, by contrast with the Globus Toolkit's exacting 'contributor's agreement', a developer seeking to contribute to Project JXTA – either through patches or by becoming a 'Committer and/or Project Owner' – can sign an agreement similar to the one required by the Apache Software Foundation. The contributors simply attest that to the best of their knowledge the code submitted is their own development work, and that they possess the authority to provide it and any related intellectual property to Project JXTA. In this case, the software development project had access to distribution capabilities, giving Sun the freedom to make use of the less restrictive and more familiar terms in the Apache Foundation's agreements.⁴³

⁴² For reasons that similarly may have been rooted in the risk averse stance of IBM and Oracle, these negotiations also resulted in the creation of a novel form of agreement governing future contributions of code to the Globus Project. Contributing developers must grant Globus a perpetual, world-wide royalty-free license to use the submitted code, and automatically would lose their right to use Globus Toolkit were they to file suit for infringement of their intellectual property rights in their contribution(s). See David and Spence (2003/2004).

⁴³ It may be noted that Sun sells a range of complementary JXTA products (software applications) that are available under proprietary software licenses, as well as commercial software systems services. On the emergence of business models that build commercial offerings around free and open source software, see the survey in S Arora and P A David, 'Commercialization of Open-source: Symbiotic or Parasitical?' (SIEPR-Project NOSTRA Working Paper, Stanford University, May 2003 - papers from this project are available at: <http://siepr.stanford.edu/programs/OpenSoftware_David/NSFOSF_Publications.html>); Carlo Daffara, *Business Models in FLOSS-based Software Companies* <<http://opensource.mit.edu/papers/OSSEMP07-daffara.pdf>>.

A fifth set of complexities may arise because the parties to collaboration, or their institutional employers, are in several different legal jurisdictions. In such a circumstance, one or more of the relevant parties may assert that issues concerning the existence, interpretation and enforcement of a purported contractual relationship should be determined according to the law of their own jurisdiction or by their own courts. They may also assert that issues concerning the first ownership of the fruits of collaboration are matters for their own law and their own courts. This becomes a matter for the complex rules as to the conflict of laws.

Questions concerning the law applicable to disputes concerning collaboration contracts will be governed in the English courts by the terms of the 1980 *Rome Convention on the Law Applicable to Contractual Obligations*. Under the terms of this Convention, the parties can choose the law that will govern disputes as to their contractual relationship, including its validity, by carefully worked out choice of law clauses included in express agreements. Yet, the ability of the parties to choose the law applicable to contractual disputes concerning intellectual property sometimes will be limited under the Convention by the application of a 'mandatory' rule as to the applicable law. In situations where no such choice of law clause is included, complex rules will determine which law ought to govern the contract. Under Article 4 of the *Rome Convention*, this will be the law of the country which has the closest connection with which the contract or, in cases concerning the existence of a contractual relationship, with the purported contract. In ordinary commercial situations, however, a number of more or less clearly formulated presumptions govern the issue of what constitutes connection with a particular country. But, it would be very difficult to predict what might be taken to be 'connected with a particular country' in the context of a collaboration agreement that pertained to activities carried on in several different jurisdictions, none of which could be said to be the 'home' of the project. Legal doctrine concerned with 'conflict of laws' has yet to come to terms with activities that take place in no particular territory. Once again, it should be emphasised that issues concerning contracts that involve intellectual property rights pose special legal difficulties.

Questions as to the jurisdiction of the English courts in disputes concerning collaboration contracts are likely to be governed by a regulation of the European Union, *Council Regulation (EC) 44/2001 of 22*

December 2000 on jurisdiction and the recognition and enforcement of judgments in civil and commercial matters. Article 23 of that Regulation permits parties to make agreements as to the courts that will have jurisdiction over their disputes. This provision applies when one of the parties is domiciled in a Member State of the European Union, but a similar possibility exists at common law. In the absence of such an agreement, extremely complex rules will determine jurisdiction, these rules starting from the presumption that a defendant ought to be sued in his or her own jurisdiction. The rules governing jurisdiction in the absence of express agreement will again vary depending upon whether the defendant is, or is not, domiciled in a Member State of the European Union. Moreover, disputes involving multiple parties from different countries add a further layer of complexity to the rules determining jurisdiction. It is essential that the arrangements underpinning international collaborations include clear and effective choices about both the laws applicable to the collaboration and the courts in which disputes concerning it will be heard.

A sixth set of problems arises because of the expertise required to interpret contracts for scientific collaboration. It may be that the parties to a scientific collaboration cannot agree on the courts in which they would like their disputes to be resolved simply because they do not believe that any court in any jurisdiction has the expertise to interpret the contract. This may either be because the contract touches upon technical issues, or because they want the contract to be interpreted by someone who understands the cultural norms of the scientific community in which they are operating. In such circumstances, the parties may enter into an agreement that their disputes will be subject to arbitration or mediation rather than to the jurisdiction of a court. In most jurisdictions there is provision, as there is in the United Kingdom under the *Arbitration Act 1996*, to render the decisions of arbitrators binding in most cases and the parties may well choose to do so in situations in which they have chosen an arbitrator for her scientific expertise. In such circumstances, the parties may also agree to the legal or other standards by which their dispute is to be resolved. Arbitration can be relatively less expensive than court proceedings and the choice of an arbitrator to resolve disputes may solve some of the difficulties associated with international collaborations.

From these doctrinal considerations, then, it is clear that close attention must be paid to the legal basis of a collaborative relationship right at the outset of the project. But the potentially high cost to the project of negotiating its every possible outcome is extremely high. The third part of this chapter is devoted to developing a mechanism for minimising that particular problem. To anticipate, it seems that there is a need to develop an array of standard contractual clauses, covering particular issues that are likely to arise in e-Science collaborations and a set of principles for the development of such agreements. This approach is not to advocate the introduction of a set of standard e-Science collaboration agreements. The formulation of such agreements would be an almost impossible task given the likely diversity of e-Science projects and the fact that the uses of the technology supporting on-line collaboration are currently largely unknown. Standard form agreements may petrify norms at an early stage in the development of the social practice of e-Science. Moreover, standard form agreements are fraught with legal risk. The development of standard clauses, rather than standard form contracts, and principles for the use of those clauses is likely to yield a far more reliable mechanism for producing robust collaboration agreements at a minimum cost to the relationship of those involved.

2.3.2 Issues Regarding Material Contributed to Collaborations

Three questions arise in relation to the material that is contributed to on-line collaborative projects. The first relates to the legal risk to which a party to such collaboration might be exposed by the use of material contributed to it by another participating party. The second relates to the extent to which the pooling of resources protected by intellectual property rights might constitute a breach of the rules of competition law. A third relates to the extent to which the contribution of particular resources gives a party a claim to the product of a particular collaborative research activity.

2.3.2.1 Legal Risk

Take first the question of legal risk. An important feature of establishing trust between the partners to on-line collaborative research will be the extent to which they can confidently use the resources that each brings to the collaboration. Each party will need to know that her involvement with the project has not exposed his or her to unanticipated legal risk to

third parties. Assume that two parties to a collaboration bring material together for a database in the health care sector. The material that they bring may well be the subject of data protection rights, rights against breach of confidence, or intellectual property rights such as copyright and database right, which are violated by its use as part of the project. A party to a collaboration who receives and uses such material may be in breach of her obligations under data protection or intellectual property laws – even though ignorant of the fact that the data in question contains material for whose use adequate consent was not obtained, and thus constituted material infringing an intellectual property right. In general, information received in breach of an obligation of confidence will be protected only if there is notice at some time before the information is used of the circumstances in which the information has been obtained. (The requisite notice, however, may be constructive rather than actual.)

At least in relation to obligations as to data protection, copyright and the database right, this risk may be less great than it at first appears. Under the terms of the European law of data protection – enacted in the United Kingdom as the *Data Protection Act 1988* – consent will not be necessary for the secondary use if the data does not relate to identifiable individuals. It will, in any case, be available to scientific research without additional consent, as long as it is not being used to support decisions with respect to particular individuals, or used in a way that is likely to cause substantial distress to an individual data subject. Similarly, an exception to infringement of copyright and database rights exists as long as the use of the protected material is for research the use is not ‘commercial’.⁴⁴ Importantly, however, no broad research exemption exists in relation to patent rights. Although patent lawyers often talk of a research exemption, in practice it is extremely limited.⁴⁵ The risk that legal consequences attach to the use of material brought by a partner to

⁴⁴ The reference here is to the *Copyright and Rights in Databases Regulations 1997* section 20 and the *Copyright, Designs and Patents Act 1988* section 29. The latter is soon to be amended under the terms of Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 – on the harmonisation of certain aspects of copyright and related rights in the information society.

⁴⁵ For the UK see *Monsanto Co v Stauffer Chemical Co* [1985] RPC 515 and *Smith Kline & French Laboratories Limited v Evans Medical Limited* [1989] FSR 513. For the US, see *Maday v Duke University* 307 F3d 1351 (2002), *cert denied* 539 US 958 (2003).

collaboration is a real one. This risk is exacerbated in the context of the use of the potentially unlawful material in international collaborations, and particularly those that take place over the Internet. But even in the straightforward situation involving the law of only one jurisdiction, the problem of legal risk remains an important consideration in the establishment of collaborations.

The lawyer's answer to the foregoing problem is that the parties to collaboration ought to take cross-indemnities from one another as regards liability arising from the research material that each expects to provide. There are two difficulties with such a solution. The first, though somewhat theoretical, is that it will not protect a party from criminal responsibility in the unusual situation in which use of the material constitutes an offence in some jurisdiction whose courts find its law applicable. The second, and more real, difficulty is that the negotiation of an indemnity clause may operate to undermine, rather than to build, the trust that is necessary to establish an effective working relationship. This effect may be exacerbated when, as is usual, it is an institution that is entering a collaboration agreement and not individual collaborators. In most circumstances it will be an employing institution that will be vicariously liable for the wrongful acts committed by those in its employment and lawyers for the institution may well seek the strongest protection that can be negotiated. This is another situation in which the development of standard form clauses and principles for their implementation may effectively avoid the potential costs to a collaborative relationship of the need to anticipate the problems that a project might entail before it has begun. Once such indemnities became established in a field of scientific collaboration, the inclusion of them in a contract governing a particular relationship betokens less suspicion of a particular collaborator.

Of course, another and more positive way of addressing these questions about the legality of the materials used by a collaborative project is to see it not as a problem of legal risk, but as a matter of good information management. The United Kingdom e-Science 'Pilot' project known as CLEF is working to build a system that connects databases in the healthcare sector. The framework will integrate clinical histories, radiology and pathology reports, annotations on genomic and image databases, technical literature and web-based resources to serve the needs of patients, their families and carers, clinical professionals and

biomedical scientists, healthcare enterprises and the public at large. CLEF is acutely aware of the importance of ethical and legal issues regarding, in particular confidentiality and intellectual property. One of the stated goals of the project is to devise agreed policies on information governance and technical measures for their enforcement. CLEF follows a two-pronged strategy to ensure that it has permission for the use of data protected by confidentiality and intellectual property rules. The first strategy is to enlist the intellectual property owners and confidentiality watch-dogs as collaborators or supporters. The second is to foster consensus on information governance policies among these partners. The emergence of standard form contracts in the formation of collaborative databases could help to spread and to entrench these agreed standards of good information management in similar applications contexts.

2.3.2.2 Competition Law

The second issue that arises in relation to the material contributed to an e-Science collaboration concerns the rules of competition law. The parties to a contract governing a collaborative project need to be careful that their agreements are not subject to control by the relevant competition law authorities. This will occur in situations in which collaborative agreements are either collusive or constitute an abuse of a dominant position. In the European Union Articles 81 and 82 of the *EC Treaty* render such behaviour unlawful. The danger of falling foul of Article 81 arises particularly when parties to a collaborative agreement pool resources such as intellectual property in a way that has the effect of excluding competition. Imagine, for example, that a number of research groups own database rights in a series of databases which, if pooled, would effectively cover the field of a particular area of research: in some circumstances the very creation of such a pool may be anti-competitive under the principles of European competition law. Alternatively, the creation of the pool may be allowed either under the terms of Article 81(3), which exempts from the operation of Article 81 certain agreements relating to research and development and the licensing of its results; it also could be found permissible under the terms of the technology transfer block exemption.⁴⁶

⁴⁶ *Commission Regulation (EC) No 240/96 of 31 January 1996* on the application of Article 85(3) of the Treaty to certain categories of technology transfer agreements.

In some other circumstances the resources to which a collaborative project gives rise may also create competition law problems under the principles expressed in cases such as *C-76/89R*, *C- 77/89R* and *C- 91/89R RTE and Others v Commission* [1989] ECR 1141. Such a situation would occur when the fruit of a collaborative research project becomes established as a dominant industry standard in a particular area of technology. The Globus Toolkit example discussed below offers an instance of just such a suite of technical specifications which may emerge as the *de facto* Grid standard. But there the issues of potential abuse of monopoly rights over an ‘essential facility’ have, in a sense been fully anticipated by the Globus Project, and therefore are unlikely to arise; the Globus Project and its industrial partners have agreed that GT is to be distributed as ‘open source’ software package on a royalty free basis, under the terms of a (newly designed) form of public license.

In this connection it may be remarked that the use of ‘free and open source software’ (F/OSS) licenses such as the GNU GPL obviously offers an attractive way of invoking copyright law to deal with many of the difficult legal complications that otherwise arise from the ownership of intellectual property arising from ‘horizontal’ collaborations (i.e., collaborations at the same stage of production and distribution). But, F/OSS licenses equally can facilitate the resolution of issues arising among the parties to ‘vertical’ collaboration agreements – as has been demonstrated in the case of the Globus Toolkit. Where public funds support the development of infrastructure technologies of this sort, especially network technologies for which interoperability standardisation is critical, such solutions might be fully justified on economic efficiency grounds. Yet, at present the United Kingdom and other governments have held back from embracing this approach as a general principle. Indeed, as has been noted, the recent direction of government policies vis-à-vis PRO’s has been to encourage the generation of revenue through the use of proprietary software licenses.⁴⁷

⁴⁷ Copyright law has been invoked by commercial software vendors in distributing programs physically and electronically under a variety of ‘shrink-wrap’, ‘click-wrap’ and ‘browse-wrap’ licenses. In David and Spence (2003/2004: Appendix 4: section 2.3), notice is taken of an objection to the royalty-free terms of the Globus Toolkit Public License made by an evaluator of the GT protocol suite: the commentator, from a British university pointed out that the UK government, and the researchers themselves, expected to see their middleware (platforms) and applications generate revenues from licensing agreements.

In this environment the issues of claims to the fruits of ‘essential’ research tools created by publicly supported collaborative projects will have to be thought through very carefully with regard to the implications of competition law as well as intellectual property law.

The whole issue of the potential competition law problems involved in the establishment of e-Science collaborations and the exploitation of their fruits has been thoroughly addressed in the recent European Research Area Expert Group *Report on Strategic Use and Adaptation of Intellectual Property Rights Systems in Information and Communications Technologies-based Research*. The conclusion of that Report is that the current application of the law to collaborative research remains unclear, and that:

... some clearer guidance from the Commission would be welcome on how it sees competition law applying in situations when industry standards require the use of a technology that is IP-protected and when access to research tools that are IP-protected is denied, or granted only at unreasonable rates.⁴⁸

2.3.2.3 Individual contributor's claims to the fruits of collaboration

The third question to be considered regarding materials contributed to research collaboration is whether one party's contribution can give rise to a claim to the fruits of the whole project. This is a very complex legal matter. The nature of the answer turns, in large part, upon the nature of the contribution. Whether or not the latter will give rise to such a claim depends largely upon the rules as to the ownership of the fruits of a collaboration – a subject to be considered in the section following this. This is an important matter that deserves emphasis because when the parties to a collaborative project consider the allocation of the fruits of the collective efforts, they quite naturally tend to focus on the resources that each has brought to the project. But, as will be evident from the discussion in the following section, the significance that the parties attach to the nature (and scientific or economic value) of their respective contributions is not reflected in the legal rules concerning the allocation of first ownership of the fruits of a (scientific) collaboration.

⁴⁸ European Research Area Expert Group *Report on Strategic Use and adaptation of Intellectual Property Rights Systems in Information and Communications Technologies-based Research* (2003) 26.

2.3.3 *Issues Surrounding Material Arising from Collaboration*

A scientific collaboration may give rise to ‘information goods’ and ‘knowledge resources’ of many different kinds. These ‘outputs’ may be protected as confidential information as long as they are kept secret. In some jurisdictions they may also, though very rarely, be protected by the law of tort under the general action against misappropriation. This tort prevents one party from ‘reaping without sowing’ by exploiting a valuable intangible that another has created and, in effect, amounts to an uncodified system of intellectual property protection. However, this tort is unknown in many jurisdictions, including England, and is very limited in its application in those in which it is recognised. In most circumstances, the fruits of collaborative research will be protected, if at all, only by the statutory intellectual property codes. Thus the fruits of a collaboration will be able to be captured if it constitutes, for example, an invention that can be patented; a work (including a computer programme) that is the subject of copyright protection; a database over which the database right can be asserted or a plant variety which can be protected by plant variety rights.

A significant level of international harmonisation of the intellectual property codes has taken place in the last 150 years. This occurred first under a series of multi-lateral international treaties that emerged in the 19th century and has more recently come about under the *Agreement on Trade Related Aspects of Intellectual property Rights including Trade in Counterfeited Goods* (the so-called *TRIPS Agreement*).⁴⁹ Nevertheless, the rules concerning the ownership of intellectual property are significantly different in different jurisdictions. The discussion here will call attention to some interesting and important differences between the relevant bodies of law in several national jurisdictions, but the determination of which nation’s law is to govern, and what court will hear disputes when the latter involve international collaborators, again is a very complex matter that must remain beyond the bounds of this review.

In the case of intellectual property disputes, the applicable law usually will be the law under which protection is sought, although the application of this principle is far from straightforward. In relation to

⁴⁹ The *TRIPS Agreement* is an appendix to the *Marrakesh Agreement* establishing the World Trade Organization (1994).

the forum in which such a dispute should be heard, Article 22(4) of *Council Regulation (EC) 44/2001 of 22 December 2000* on jurisdiction and the recognition and enforcement of judgements in civil and commercial matters and equivalent common law rules, provide that questions concerning the registration or validity of an intellectual property right will be heard in the courts of the jurisdiction in which the intellectual property right is claimed. So, for example, questions as to the validity or registration of a German patent will need to be heard in the Bundespatentgericht. By contrast, however, the conflict of law rules of many countries, including those of the United Kingdom, provide that infringement actions can be heard in the courts of other jurisdictions, often those in which the defendant is domiciled.⁵⁰ The context in which an alleged infringement takes place over the Internet raises even more uncertain questions of jurisdiction, particularly in light of the growing tendency of the courts, evinced in cases such as *Menashe Business Mercantile Ltd v William Hill Organisation Ltd*⁵¹ to claim jurisdiction over the infringement of intellectual property rights on the Internet involving activities that might be seen as taking place in another country.⁵²

Again, the question of conflict of laws presents a minefield for potential international collaborators who may find themselves either defending suits, or bringing suits to protect their intellectual property, in foreign jurisdictions and under foreign law. This cannot be entirely avoided by carefully drafted contractual provisions as to the choice of law and choice of courts, but such provisions can significantly reduce the risks to a successful collaboration presented by the arcane rules as to the conflict of laws.

The first issue to be addressed in determining whether the outcome of a collaborative project gives rise to intellectual property that can be owned is whether it falls within the subject matter requirements of the statutory regimes. The most likely outcomes of such a project will be a patentable invention, a copyright computer programme or a database which might be the subject of either copyright or database right. A consideration of the issue of the copyright and database protection of scientific databases will reveal something of the intricacy of this class of issues.

⁵⁰ See *Pearce v Ove Arup Partnership Ltd and Others* [1997] 2 WLR 779.

⁵¹ [2003] RPC 31.

⁵² See *Menashe Business Mercantile Ltd v William Hill Organisation Ltd* [2003] RPC 31.

Scientific and technical databases will be protected by copyright if they amount to original literary works. Under section 3A(2) of the *Copyright, Designs and Patents Act 1988*, a database will constitute such a work only where 'by reason of the selection or arrangement of the contents of the database the database constitutes the author's own intellectual creation.' It is therefore the manner in which the database is organised that is protected by copyright rather than its contents. There is academic debate about whether the standard of originality that must be met for copyright in a database is higher than that applied to literary works generally in relation to which is usually held that 'original' simply means 'not-copied'.⁵³ In the implementation of the *Software Directive (1991)*⁵⁴ and the *Term Directive (1993)*,⁵⁵ the United Kingdom government seems to have assumed that it did not, as it was not thought necessary to amend the law of the United Kingdom so as to meet similar provisions defining originality. By the time of the implementation of the *Database Directive (1996)*⁵⁶ the government may have changed its position, as the *Copyright, Designs and Patents Act 1988* was amended to include the same definition of originality for databases.

In the United States, the arrangement of a database may also be protected by copyright, but again the standard of originality required for protection is uncertain. The US Supreme Court's ruling in the leading case, *Feist Publications v Rural Telephone*⁵⁷ made it clear that originality must consist of more than mere effort in the compilation of information. But inasmuch as the arrangement of the information in that particular case was alphabetical, it is difficult to determine just how much originality was required and whether the standard was significantly higher than that required in the United Kingdom. The situation is clearer in the European Union, where a database might also attract the protection of a database right under the terms of the *Database Directive*, which was implemented in the United Kingdom as the *Copyright and Rights in Databases Regulations 1997*.

⁵³ *University of London Press v University Tutorial Press* [1916] 2 Ch 601.

⁵⁴ *Council Directive 91/250/EEC* on the legal protection of computer programs, art 1(3).

⁵⁵ *Council Directive 93/98/EEC* harmonising the term of copyright protection, art 6.

⁵⁶ *Directive 96/9/EC* of the European Parliament on the legal protection of databases, art 3(1).

⁵⁷ 111 SCt 1282 (1991).

This protection prevents not merely the reproduction of the selection or arrangement of a database, but the extraction or re-utilisation of a substantial part of the contents of the database. In order for a database to qualify for this protection there must have been substantial investment in obtaining, verifying or presenting the contents of the database. Copyright protection in the United Kingdom will be available to nationals of most countries. By contrast, the Database right will be available only to nationals of the European Economic Area, although it will also be available to someone not from the European Economic Area who jointly makes the database with a person (or company) who is resident there. Despite the inducement for reciprocal introduction outside the EU, which was seen to be the motivation for this departure from the ‘national treatment’ provisions familiar under the *Berne Convention* on copyright protection, attempts to introduce similar legislation in the United States have so far not met with success.

A problem arises under the database right because of the content of many databases, and, indeed, most of the particularly valuable scientific and engineering databases are continually changing. The UK *Copyright and Databases Regulations 1997* provide that:

17(3) Any substantial change to the contents of a database, including a substantial change resulting from the accumulation of successive additions, deletions or alterations, which would result in the database being considered to be a substantial new investment shall qualify the database resulting from that investment for its own term of protection.

This paragraph was the subject of a reference to the European Court of Justice from the English Court of Appeal in the case *C 203/02 British Horseracing Board Limited v William Hill Organisation Limited*.⁵⁸ In the court of first instance, Laddie J held that the *Regulations* were intended to protect dynamic databases that are constantly being updated. He held that as a database is updated it is subject to a new term of protection on an on-going basis but that an unauthorised user who takes older data ‘only faces a database right which runs from the date when all of that older data was present in the database at the same time.’ Moreover, if someone ‘takes an existing database and adds significantly to it, he

⁵⁸ [2004] RPC 13.

obtains protection for the database incorporating his addition.⁵⁹ The European Court of Justice found that it did not have to address the correctness of this position, but at some point it will need to be addressed.

It is clear that this legislation will be difficult to apply in many circumstances involving on-line scientific collaboration. A database may be being constantly updated. It may therefore be very difficult to determine either the point at which it has changed identity and its term of protection thereby extended or the point at which new rights in the database have been acquired by someone adding to it. Good examples of such a database might be: (i) the Comb-e-Chem project in which results from an on-line test-bed in combinatorial chemistry are stored as a part of a database that is constantly being updated by each new user, or (ii) the GENIE project which creates a constantly updated database of results arising from the use of a Grid enabled integrated earth system model.

The determination of the ownership of the knowledge resources arising from on-line collaborations is often far from simple. Assuming that the outcomes of on-line scientific collaboration include subject matter that can be protected under the various statutory intellectual property codes, it then becomes a matter of determining who has ownership over that subject matter, either under the statutory regime itself or some legally effective agreement. Most of those involved in on-line science will be working for a university or other research institution. They may also be working in partnership with a private firm or with funding from a public or private body. Importantly, they may well change employers during the life of a particular collaborative research project. Disputes will inevitably arise as to the ownership of valuable knowledge resources between these different parties.

The allocation of the first ownership of intellectual property varies slightly from jurisdiction to jurisdiction and from statutory regime to statutory regime. In general terms, however, the law of the United Kingdom allocates first ownership to the creator of the particular resource or to her employer. For these purposes, an employer is not simply someone who provides the funding for, or commissions, the

⁵⁹ *British Horseracing Board Limited v William Hill Organisation Limited* [2001] RPC 31.

research. Rather, it is someone with whom the creator is in a 'master-servant' relationship. In order for first ownership of the resource to be allocated to the employer, the material in question must be created in the course of employment. In the field of inventions, the allocation of first ownership to the employer is somewhat offset by section 40 of the *Patents Act 1977*, which provides that an employee inventor can be remunerated in circumstances in which his or her invention proves to be of 'outstanding benefit to her employer'. But claims under this provision are rare.

The application of all this law to the position of academics in the United Kingdom has been much debated. For some while it was the received wisdom that copyright in works created during university employment first vested in the academic author, whereas rights to patentable inventions first vested in the university.⁶⁰ In the wake of the decision in *Greater Glasgow Health Board's Application*,⁶¹ however, it is now widely believed that rights to patentable inventions also will first vest in an academic inventor. The continental European and American intellectual property regimes are even more complicated in the allocation of first ownership than are the rules in the United Kingdom; and some countries, such as Germany and Sweden have special rules that apply to university employees.

In all jurisdictions, these rules as to first ownership become even more complicated in situations of collaborative creation. This is particularly difficult for international collaborations because the rules as to joint creation vary more from jurisdiction to jurisdiction than do some other types of intellectual property rules. Take, for example, the rules as to copyright. In the United Kingdom, copyright will be jointly owned if the contribution of one author is not distinct from that of the others under section 10(1) of the *Copyright, Designs and Patents Act 1988*. In Germany, however, there is a requirement of joint creation and a requirement that each of the parts is incapable of independent exploitation, but it is acceptable that the different parts of the jointly created programme have been made at different times.⁶² In the United

⁶⁰ See W R Cornish, 'Right in University Innovations: The Herchel Smith Lecture for 1991' (1992) 13 *EIPR* 15–16.

⁶¹ [1996] RPC 207.

⁶² *Buchhaltungsprogramm* BGH I ZR 47/91 (1995).

States, US *Copyright Act* §101 provides that the intention of the authors as to whether their work will be merged into inseparable or interdependent parts of a unitary whole will be determinative. In practice some jurisdictions, such as the United States, have a tendency to categorise works as jointly created, and therefore have very fully articulated rules as to the way in which different authors must deal with one another and with strangers; whereas other jurisdictions, such as the United Kingdom, have far less developed law as to joint authorship because they are less likely to categorise works as jointly authored. Thus a song is likely to be treated as a single copyright work in the US, even though one person has written its words and another its music; whereas in the UK the same material would be treated as independent literary and musical works. Particular problems of joint ownership arise when a party brings to a collaboration material to which either copyrights or database rights may already apply – especially when the ownership status is not clearly known *ex ante*.

At least in the United Kingdom, it is difficult to predict whether in such circumstances the courts would ever be prepared to treat the whole resulting outcome as the subject of a new proprietary right that is jointly owned, or simply regard the resulting product to be a composite of different items of protected subject matter. In most circumstances it is likely that the latter result would prevail, but this may not always be the case. For example, under section 17(3) of the *Copyright and Databases Regulations 1997* it is plausible that a new, jointly owned database can be made out of two existing databases. The distinction between these two interpretations might become important when a researcher changes institutions. If the first employer has rights to the database that she creates in one place and she takes that database and adds to it in the new institution, the different analyses suggested here will determine whether the original university has any claim to rights in the whole of the resultant resource or simply in those parts of it which the researcher took with her when she left her first employment.⁶³

While the rules as to the first vesting of intellectual property rights in jointly created resources are very intricate, it is important also to note

⁶³ For a cautionary tale regarding the institutionally mobile author of a scientific database, see David and Spence (2003/2004): Appendix 5 (The Effect of Uncertainty as to Ownership: A Cautionary Tale - The Attwood Experience of Database Ownership).

that express agreement, or the assignment of rights can in most jurisdictions displace the effect of all these rules as to first ownership. The issue upon which to focus then becomes that of the terms of the contractual arrangements between the individual members of a collaborative project and their employer institutions and funding bodies. The Association of University Teachers in the United Kingdom has recently conducted a survey of policies concerning the first ownership of intellectual property amongst university employers in this country and has found significantly divergent practices.⁶⁴ This policy diverges even more widely amongst universities internationally, although the trend in recent years, not surprisingly, has been for universities to claim more and more. When outside parties get involved they too may have their own requirements as to the assignment of intellectual property rights. Because the funding arrangements will vary depending upon the type of project, it is difficult to generalise about arrangements for the allocation of rights.

It seems clear, however, that distributive norms need to be established in relation to the ‘proper’ allocation of rights among researchers for projects of different kinds. These norms also should be flexible enough to respond to situations in which the identities of the parties involved in a collaboration, or their employers, have altered. This is again a situation in which widely shared principles and standard contractual clauses, interpreted by competent *fora*, could be used to reduce the uncertainties surrounding the establishment of collaborative projects. The difficulty in addressing this need is that without the assistance of an institutional mechanism for the efficient resolution of these issues at the outset of the collaborative undertaking, subsequently emerging disputes over the ‘knowledge assets’ that have been created may make it difficult for the latter to be used effectively, thereby defeating the purposes of e-Science.

2.3.4 Issues of Liability Arising from Collaboration

Most thinking about on-line collaboration has focused upon questions of ownership, either of the inputs to a collaboration or of the knowledge resources to which it gives rise. An equally important issue, however, is that of professional responsibility for the conduct of a project and

⁶⁴ This is available in the Members Only section of the web site <<http://www.aut.org.uk>>.

liability for any harm that it might cause. These are linked, but not identical, questions.

The question of professional responsibility is that of who will bear the reputational or other loss associated with the discovery that research has been conducted either incompetently or unethically. The scientific community has traditional norms controlling the damage caused by incompetent work or unethical behaviour, such as the falsification of research results. More recently those norms have been given effect formally through the policies of universities, funding bodies and journals, but they also are enforced informally through the shared understandings about the subsequent treatment of malefactors by colleagues, especially those involved in the same branch of scientific enquiry. In general terms, norms of professional responsibility traditionally have reflected, and should reflect, the structure of norms for the assignment of credit for the project's scientific achievements. That is, the senior scientist with responsibility for directing the work of the project (or the facility whose resources it uses) gets most credit when it is a success, and corresponding also ought to be the person whose reputation suffers most harm if the research is found to have been poorly conducted, or the findings have been misreported.

Parallel to this question of professional responsibility, there exist legal rules that determine legal responsibility for harm that a project may cause. These legal rules may well attribute responsibility in a way that is different to that in which the traditional norms of the scientific community would do. Moreover, questions of liability might arise even for parties who do not engage in e-Science collaboration themselves, but merely provide the platform for others to do so, such as DiscoveryNet or myGrid. In England, the law of tort often will determine the question of whether, and how, harm caused by a collaborative project ought to be compensated. This will be the case whether the relevant harm has been suffered by one of the parties to a project or by a third party. In broad terms, the question will frequently become one of whether the behaviour which has given rise to a particular harm was negligent in that the harm was reasonably foreseeable and could have been prevented. For example, the rapid spread of a new virus through a computer system (causing losses of data, or damage to instruments controlled by that system) may be deemed to have been impossible to guard against, whereas the failure to update available anti-virus filters in

a system firewall could be more readily construed as negligence. Negligence, and indeed the law of tort, is far from the only potential sources of liability to which the participants in collaborations may be exposed. For example, parties also might be liable because the project uses or creates material that is in breach of some intellectual property right, or because confidential data is accidentally disclosed, violating the privacy of outside parties.

Collaborative projects give rise to a host of potential questions of liability. As is the case with issues of responsibility for the legality of inputs to a research project, the issue of responsibility for the conduct of research and its outcome needs to be addressed by the parties to a collaborative project before it is launched. The lawyer's solution is likely to be a combination of indemnities and insurance for the most obvious risks, but this again raises the various costs of negotiating an e-Science collaboration. Once more the discussion has moved into an area of concern where it would be most helpful to promote the articulation of common principles for collaboration which reflected the traditional norms for the allocation of responsibility in the scientific community, and where the development of standard form contractual clauses consistent with those principles clearly would be recognised as desirable.

The purpose of the foregoing review has not been to offer an exhaustive list of the legal issues that will arise in the context of e-Science collaborations. Rather, it is intended to point to the range and complexity of the issues that are likely to arise. It is clear that these issues cannot be addressed in a single programme of 'law reform'. Equipping the scientific community to deal with the complexities of the institutional environments in which they work demands a more subtle and responsive institutional design of the kind that is outlined by the next Part.

3. A PROPOSED INSTITUTIONAL DEVELOPMENT PROCESS FOR e-SCIENCE

To devise one or several approaches for arriving at institutional mechanisms whose establishment would generate workable specific arrangements that facilitated collaboration in e-Science is a considerable challenge. The multiple parties and jurisdictions involved in e-Science collaborations, and the need to balance conflicting interests among

them, make the design of effective governance arrangements an extremely complicated and thorny problem. The costs of getting those arrangements even slightly wrong can be very high.⁶⁵ Furthermore, most research scientists have socially more productive things to do than spend their time thinking about how to arrive at a good set of governance mechanisms, even though it would be impossible for others to achieve that goal without thoroughly involving the affected research communities in the process. More difficult still will be the task of obtaining quickly negotiated contractual arrangements that also could be of use to facilitate collaborative activities in other spheres where Grid infrastructure is likely to be extended – including e-Learning, e-Government, e-Commerce, e-Healthcare.

It is clear that these appropriate institutional mechanisms cannot simply be put into place by legislation and that the problems created by the international nature of collaborative e-Science cannot be solved by the international harmonisation of formal legal rules. Legislation and the harmonisation of legal rules have a potentially stultifying effect on the development of appropriate institutional mechanisms in this area. When legislation is enacted and international conventions are agreed, they tend to have the effect of petrifying the norms regulating a given area of behaviour for a long period. This poses the risks that the norms which have been set may have been set at a particular point in the development of a social practice, such as on-line scientific collaboration, and may rapidly become inappropriate. The history of the international agreements on the protection of semiconductor topographies provides a good illustration of the need to proceed in a way that can avoid this potential danger.

3.1 Creating an Appropriate Institutional Mechanism: Basic Considerations

Any standardisation of the norms that govern e-Science must be sufficiently flexible to undergo non-disruptive evolution with the development of mode of organising and conducting research that is, after all, still in its infancy. Moreover, the international harmonisation of legal rules is unlikely to be effective. The international harmonisation of

⁶⁵ See David and Spence (2003/2004): Appendix 6 (The Effect of Inappropriate Licensing: A Cautionary Tale-The SWISSPROT Experience of Database Licensing).

law is a slow and frustrating process as the *TRIPS* negotiators have found. Harmonisation would be a particularly daunting task given the range of legal issues that might impact upon the conduct of collaborative on-line research. Moreover, the harmonisation of legal norms is only ever partially effective in achieving the goal that disputes determined under the same norms will find the same result in different courts. The history, for example, of the European *Patent Convention* shows that the same norms can lead to different outcomes in different courts with different interpretative traditions. Formal law reform and the harmonisation of laws do not seem to be the answer to establishing norms that can facilitate collaborative e-Science.⁶⁶

Consequently, the nub of this chapter's recommended approach to constructing appropriate institutional infrastructures for e-Science is the creation of a co-ordinating and facilitating mechanism, in the shape of a novel public agency. We envisage the establishment of an independent body that could be formally designated as the 'Advisory Board on Collaboration Agreements' (ABCA). Its remit would be to guide, oversee and disseminate the work of producing, maintaining, evaluating and updating standard contractual clauses, those being the constituent elements from which formal agreements may be more readily fashioned by the parties undertaking specific 'Grid-enabled' collaborations in science and engineering research. This advisory board would, of necessity, play a leading role in enunciating a set of fundamental principles to guide the formulation of those contractual clauses and ensure that the effects of the agreements into which they are introduced will not be inconsistent with the intent underlying those principles.

In view of the importance of finding some suitable response to the needs that have been described by the preceding discussion, some new institutional departures along these lines appears very much in order. That appropriate institutional mechanisms can make a critical

⁶⁶ This legal approach is neutral with regard to the balance struck in such agreements between providing for commercial exploitation of intellectual property rights and protecting the public domain in scientific and technical data and information. One should note, however, that it can accommodate the thrust of recent proposals that address the latter concern, most notably, J H Reichman and P F Uhler, 'A Contractually Reconstructed Research Commons for Scientific Data in a Highly Protectionist Intellectual Property Environment' (Winter/Spring 2003) 66 *Law and Contemporary Problems* 315–462, forthcoming. Explicit recommendations favouring such an approach are advanced below, in Part 4.

contribution to the success of the United Kingdom's investments in e-Science seems beyond serious dispute. There is no point in investing in new technologies to facilitate and empower complex collaborations if at the same time we are imposing rules and regulations (including legal and institutional administrative arrangements) that excessively raise the costs of actually carrying out those sorts of collaborations. So we need non-technological governance mechanisms (of which institutional regulations and legally sanctioned contractual forms constitute important examples, but not the only ones) whose effects tend to reinforce, rather than to counteract, those of the technological infrastructure of e-Science. To initiate a deliberate movement towards this goal, this section of the chapter proposes the launching of an exploratory process of institutional learning. It identifies the types of expertise that should be engaged, and the forms in which it could be mobilised by the United Kingdom agencies that would fund the suggested programme of consultative research and institutional experimentation.

3.2 An Independent 'Public Actor' – The Proposed Mechanism and Some Existing Models

What is needed is the establishment of a new 'public actor', a separate agency with on-going powers to initiate, co-ordinate and provide resources required to support and, above all, articulate principles for developing an array of model contractual clauses. Each of these clauses would address a particular problem among the myriad of legal issues that have been seen to arise from the formation of research collaborations, and variant solutions would be provided by the clauses developed under each topical heading. Much of this detailed work could be entrusted to specialised task force-like committees – possibly resembling the many 'study committees' set up under one or another of the US National Research Council Boards.

The activities of the study committees organised under the auspices of what we here call 'the ABCA' would focus upon framing appropriate standard contractual clauses that could be readily assembled into a variety of alternative collaboration agreements, much in the same way that software sub-routines and modules can be assembled into functionally more comprehensive software systems that are suited for particular applied tasks. As part of its supervisory and co-ordinating role, the ABCA would have not only to think about the underlying

principles that will be implemented through the contractual clauses of those agreements. It would also need to determine the best ways of organising the accumulation and dissemination of information and analyses concerning the actual formulations and manner of implementation of contractual agreements. These principles for the establishment of e-Science collaborations and model contractual clauses could then be put into effect in individual cases by universities and research bodies.

The part of the ABCA's activities which involved promoting principles of best practice in establishing institutional arrangements for e-Science collaborations might find a parallel in the work of the Basel Committee on Banking Supervision.⁶⁷ Under the Basel Accord the Governors of the Central Banks of thirteen countries develop common principles for banking supervision. The Basel Committee on Banking Supervision does not make laws of any kind, rather it builds consensus between important actors in the international banking community and these standards are given appropriate effect by relevant actors in local contexts. The experience of the banking community has been that this approach has the advantage of flexibility and that principles develop slowly rather than being imposed at what might be an inappropriate stage in their articulation. In time, it might be hoped that an international body for the development of collaborative research principles might be established, similar to the Basel Committee. However, the establishment of a national body with such a task would be an important step in establishing socially desirable rules for the organisation of e-Science at a national level. Indeed, it is likely that a United Kingdom body charged with the functions of the ABCA could set a lead in the organisation of e-Science collaborations and that its principles and contractual clauses would be widely adopted even by those not a part, either direct or indirect, of the United Kingdom e-Science network.

The part of the ABCA's work that involved the development of standard contractual clauses might also find a parallel in the work of the

⁶⁷ The Basel Committee is briefly considered as a model in the context of e-Commerce regulation in D Casey and J Magenau 'A Hybrid model of Self-Regulation and Governmental Regulation of Electronic Commerce' (2002) 19 *Santa Clara Computer & High Technology Law Journal* 1, 27.

Basel Committee and, with one important reservation, in that of the Grain and Feed Association ('GAFTA'). The Basel Committee makes recommendations for contractual clauses in certain areas of banking practice as a means of ensuring that the principles it articulates are given appropriate and certain effect. Central amongst the clauses that the ABCA might be expected to suggest would be those reflecting a consensus as to the appropriate *fora* for the resolution of disputes under e-Science collaboration agreements, particularly those involving parties from different jurisdictions. A model of how such choice of forum clauses might operate can be found in the standard form agreements established and maintained by GAFTA. GAFTA has 80 different standard contracts under which more than 80 million tonnes of the world's trade in cereals and 70% of trade in animal feeds moves annually. These contracts contain arbitration clauses that allow parties to make use of the GAFTA Dispute Resolution Service. This service provides for the speedy and final resolution of disputes in an expert forum that is cheaper and quicker than traditional legal systems; further, it has the advantage of being outside the legal system of any of the parties to the international collaboration. The process involves a possible appeal to a GAFTA appeal board. Awards given under the arbitration system are enforced either informally through publication of the fact that a party has failed to comply with an agreement or, as a last resource, through the court system of the jurisdiction in which the party against whom the award has been made under the *New York Convention on the Recognition and Enforcement of Foreign Arbitral Awards 1958*. In the event that arbitration fails for some reason, disputes in international commodities contracts of this type are by agreement usually referred to the courts of a small number of jurisdictions, in particular those of England and New York. The establishment of standard arrangements for the resolution of disputes in transnational e-Science collaborations would greatly reduce the uncertainty surrounding such projects.⁶⁸ It is important at this point to emphasise why the approach recommended is one of standard contractual clauses and principles for the development of collaboration agreements rather than standard form agreements of the

⁶⁸ Yet another parallel might be found in the highly effective matrix of formal and informal norms that regulate the international trade in cotton under the rules of the Liverpool Cotton Association, see L. Bernstein, 'Private Commercial Law in the Cotton Industry: creating Cooperation Through Rules, Norms and Institutions' (2001) 99 *Michigan Law Review* 1724.

type established and maintained by GAFTA. This is for three reasons. Firstly, the potential contexts of e-Science collaborations are very various and the formulation of standard form agreements would be an almost impossible task. Even GAFTA must maintain 80 agreements in relation to the relatively routine transactions that make up the international trade in grain and animal feeds. The contexts in which standard form contracts tend to be most successful are those such as domestic conveyancing in which only a very limited range of issues is likely to arise.

Secondly, the introduction of standard form agreements, like legislation and the harmonisation of laws, may have the danger of petrifying norms at an inappropriate stage in their articulation. We are at the beginning of a new era of scientific collaboration, based upon high bandwidth telecommunications and grid-enabled computing, and to put standard form agreements in place at this moment may entail the danger of ossifying the development not only of appropriate norms for e-Science, but also of inhibiting flexibility in the elaboration of the enabling technological infrastructure. The absence of extensive experience (and hence of the weight of precedents) concerning arrangements in the new environment creates an opportunity to exercise greater leverage over the future evolution by setting standards early and firmly.⁶⁹ Yet, the very same conditions make it difficult to gauge what the new standards should be.

The fact that principles, more obviously than standard form agreements, must be developed over time will serve to emphasise to the community of lawyers and administrators in academic institutions and funding bodies that the agreements they produce need to reflect changing scientific practice.

⁶⁹ This is the non-technical form of the classic problem of 'anticipatory standard-setting' in regard to network interoperability standards, sometimes referred to as 'the Blind Giant's problem'. See P A David, 'Some New Standards for the Economics of Standardization in the Information Age' in P Dasgupta and P L Stoneman (eds), *Economics and Technological Performance* (1987) 206–39; P A David, 'Standardization Policies for Network Technologies: The Flux Between Freedom and Order Revisited' in R Hawkins, R Mansell and J Skea (eds), *Standards, Innovation and Competitiveness: The Political Economy of Standards in Natural and Technological Environments* (1995) 15–35; P A David and M Shurmer, 'Formal Standards-Setting for Global Telecommunication and Information Services' (December 1996) 20(10) *Telecommunications Policy* 789–815.

Thirdly, the blanket use of standard form agreements is fraught with legal risk. This is because such agreements can come to be used without appropriate legal advice and institutional support. Indeed, this is more likely in the context of e-Science than it is in context of international trade. Three particular legal problems tend to emerge when standard form agreements are used without appropriate professional advice. In the first place, it is obvious that a standard form agreement may not be appropriately adapted to the project which it is intended to govern. Second, in situations in which the parties simply adopt a standard form without addressing their minds to the appropriateness of its terms, the courts and legislators in most jurisdictions have come to regard them as inherently suspect. This is because contractual liability is generally, though not universally, regarded as essentially consensual. In the case of a standard form agreement, it is sometimes difficult to see how at least one of the parties can be taken to have consented to all its terms. The issue of how consent can be given to standard form agreements on-line has been thought to give rise to particular problems that have attracted both legislative and academic attention around the world. Third, when different standard forms are used by the parties entering a particular transaction, each can purport to rely upon her own usual form; it then may be difficult to determine which, if indeed either of the two had ultimately been adopted as the basis of the agreement. This is sometimes referred to as the 'battle of forms' and will become a problem as 'form shopping' encourages a proliferation of standard form agreements.

It is therefore suggested that the approach of developing standard contractual clauses covering most of the issues likely to arise in establishing and conducting an e-Science project, together with principles for their use to be observed by lawyers and administrators in academic institutions and funding bodies is a better way of developing this area of law than the development of even a stable of alternative standard form contracts. What is called for, in effect, is the institutionalisation of an iterative, adaptive procedure for developing 'meta-agreements' – contractual analogues of 'meta-standards' in the technological domain. This way of navigating the legal thicket can be likened to a technological meta-standards approach based upon setting 'performance standards' rather than technical specifications, which has long been advocated as most appropriate for anticipatory technical

standards development purposes. (In the latter, almost as a matter of definition, the relevant underlying technology and the markets are evolving rapidly; this means that very substantial uncertainty surrounds the capabilities of new systems and the attributes that their users will most value.)

Setting performance standards, however, implies that criteria of acceptability must be defined in the relevant dimensions, and procedures for compliance testing and certification need to be established. Although those sometimes are held to be impracticably costly requirements to impose in the technological sphere, an analogous process would appear to be considerably less problematic (indeed, quite natural) when applied in the domain of contractual clauses. The apparatus already exists for the effects of the latter to be examined by reference to expert opinion about pertinent legal doctrine, the empirical experience of contractual negotiations, and the law courts' rulings in a variety of jurisdictions.

In summary, therefore, three key conditions seem necessary for the ABCA to be effective in its undertakings. First it is necessary that it delineate readily intelligible transcendent principles for agreements governing collaborative e-Science projects – across as wide a variety of academic research domains as is possible. Second, to be workable and substantially self-enforcing, these guiding principles also must reflect the essential values of the scientific communities whose collaborative work is to be facilitated by institutional instruments (and the technological infrastructures of Grid-enabled computing). Third, in order to yield contractual agreements that are flexible enough to accommodate distinctive features of the relevant research community norms, as well as of the requirements of the particular science that is being planned in each case, the ABCA should eschew trying to write 'model contract' or 'standard form' agreements, and focus instead upon the development of standard contractual clauses.

3.3 Organisational Form and Composition of 'The ABCA'

The precise form of the responsible independent body thus envisaged can best be left for discussion and subsequent determination, once there is substantial agreement about its goals and operating procedures. One approach deserving serious consideration would be to have the

proposed advisory board on collaboration agreements would independently from, but in close liaison with bodies representing the pertinent administrative agencies – including the office of the Director of the Research Councils, the Higher Education Funding Council for England (HEFCE), the Scottish Funding Councils for Further and Higher Education, Universities UK (UUK) and the Standing Committee of Principals (SCOP).

Alternatively, the establishment of the ABCA could be approached more gradually, by organising a continuing ‘Working Party on Agreements’ which reported regularly to the Joint Information Systems Committee (JISC-WPA). Once a body of contractual clauses and information about the circumstances in which they proved most suitable had been developed, the Working Party might be dissolved and replaced by the ABCA envisaged above. Eventually, as its recommended contractual clauses came to be more widely used for publicly funded collaborative projects in e-Science, the ABCA might eventually be transformed into a completely free-standing and permanent official government ‘Commission on Institutional Infrastructures for Collaborations in e-Science’ (CIIC:e-Science). The colon in the latter looks ahead to the formation of a succession of such commissions, each having analogous responsibilities for assisting the formation of collaborative agreements in a different sphere, e.g., e-Commerce.

The composition of the ABCA’s appointed study committees, like the membership of that body itself, would have to be multidisciplinary. The working groups particularly need to draw upon technical expertise regarding the hardware and software infrastructures supporting collaborative e-Science, and the complex systems of resource allocation of which scientific and technological research and teaching is a part. Moreover, through the guidelines provided by the ABCA as part of the committee’s remit, it must be attentive to the larger societal goals and values that HEIs and public research institutes need to serve. The e-Science Board’s members also need to stay informed and especially foresighted regarding potential opportunities and challenges that are likely emerge when the technological infrastructures created for e-Science are opened for the development of applications supporting e-Learning, e-Commerce, e-Government, and so on.

In embarking on what must be an evolutionary programme of research, practical experimentation, assessment and redesign, it will be vital to

enlist the assistance of lawyers (both academic and practising) who are well versed in the array of legal issues identified by this chapter – and with their treatment not only in common law, but under other legal systems. No less important for this work will be the recruiting of a core cadre of social scientists who have been involved in science and technology policy studies in the United Kingdom and other leading research countries. They should have particular expertise in the social and economic organisation of academic research communities, as well as with corporate research management practices and policies regarding intellectual property rights. They should be complemented by rotating groups of experienced practitioners drawn from two constituencies: senior scientists familiar with the variety of research communities that are likely to take the lead in moving their activities onto e-Science platforms, and representatives of university administrations who are engaged in solving the practical problems posed by research collaborations.

3.4 Requirements for Information on the Experiences of e-Science Collaborations

The development of a clearer picture of the institutional context of collaborative e-Science therefore can be viewed as one of the derivative implications of the recommendations advanced by this chapter. Evidently, this entire proposed programme of legal mechanism design will need to be informed by systematic data collection about the informal ways in which disputes among collaborating researchers, and among institutional partners too, actually may be resolved before the parties enter litigation. Corresponding research is necessary about the circumstances in which scientific and technological projects are most frequently delayed, or irremediably disrupted by conflicts involving contractual matters. A multidisciplinary inquiry into the role played by institutional infrastructure factors in the experience of successful (and unsuccessful) e-Science collaborations should be initiated in conjunction with the assessment work that the ABCA should plan to carry out in regard to the effects of its own work.

The discussion (in section 2.1, above) of the current institutional context of scientific collaboration brought out the comparative paucity of concrete empirical knowledge concerning the specifics of current individual researchers' experiences with informal governance

arrangements for scientific research collaborations, and a parallel lack of systematic information about the ways that formal legal rules are being utilised by co-operating institutions. As this kind of information that could be gathered is a part of the process recommended by this chapter it is appropriate to take fuller note of this requirement; and equally of the opportunities that would thereby be opened for social science and legal studies of the changes occurring in the social and economic organisation of contemporary science and engineering research.

First, more information is needed about current practices in planning the institutional structures of e-Science collaborations. In particular, it is important to know the extent to which the working scientists themselves have an input into planning the institutional aspects of those projects, rather than simply specifying the technical requirements. If they only, or even principally, address the technical questions, there is a risk that planning will be on the basis of a 'perfect team' assumption – with issues of imperfect team behaviour therefore being left unaddressed. Further, if the actual collaborating agents have little or no role in planning the institutional arrangements, there is the real danger that those arrangements will reflect choices that are at odds with the culture of collaborative research and inimical to its success.

A further question regarding the planning phase of collaborative work is the issue of how, and from whom, the parties to collaborations, particularly those in publicly funded institutions, receive legal advice. Work needs to be done to discover the extent to which this advice is enabling, and the ways in which it creates possibilities for the establishment of individual collaborative projects, rather than simply imposes costs upon them. Empirical studies should be directed to document the extent to which existing collaborations can and do support their institutional arrangements by implementing technical controls upon access, modification and reproduction of data and information. It is as yet quite unclear the extent to which parties to scientific collaborations in academic institutions are using technical measures to monitor and enforce compliance of members with the rules of participation in a project, although employment of such devices by commercial database providers is well documented. Finally, under this head, it is important to learn whether the rules for the administration of a particular collaboration are simply agreed rules of behaviour, or

whether they are built into the very way in which the project is structured.

Second, more needs to be known about the way in which the participants in e-Science collaborations currently use the legal rules that touch upon their projects. This involves a number of issues. For example, when a dispute between collaborators arises, in what way do collaborators invoke the traditional informal norms of their particular scientific community and in what ways do they invoke formal legal rules? Evidence from contractual disputes in the commercial world suggests that formal rules are invoked only rarely, and when they are, it occurs at a point where the relationship among the parties already has become very strained.⁷⁰ It is important to know whether this pattern also holds true in the world of collaborative e-Science. Similarly, there is a question regarding the extent to which collaborators allow their relationships to develop as they proceed and the formality or informality with which they vary the agreements supporting their dealings with one another. The experience of commercial lawyers is that even parties with access to sophisticated legal advice can allow a significant divergence to emerge between the formal legal basis of their relationship and the rules upon which it actually operates. This can have unfortunate legal consequences, but it would be surprising if it were not the case in collaborative science.

More generally, there is a need to learn much more than presently is known about the interplay among the technical, social and institutional constraints on e-Science. For one thing, such studies would be of great help in validating a number of the assumptions underpinning the analysis and recommendations of this chapter, and could be expected to contribute significantly to the future design of institutional arrangements that would more effectively promote 'e-collaborations'. The sort of advisory body whose creation has been recommended here could be charged with the responsibility not only to propose effective legal devices and organisational procedures to facilitate collaborative e-Science, but also to commission on-going social science research that should underlie its work.

⁷⁰ And quite often, not even then. In many instances formal legal action is initiated only when one of the parties reaches a 'threat point' imposed by the potential for third party action (eg shareholder lawsuits or bankruptcy proceedings).

4. CONCLUSION: IPR POLICY REFORMS AND THE WAY FORWARD

In setting out the scope of the challenge of providing workable institutional conditions for productive e-Science, and in proposing a particular approach towards that goal, the preceding parts of this chapter have avoided explicit discussion of the underlying policy positions that might be reflected in the ‘working principles’ adopted by the proposed Advisory Board. It has been noted, of course, that these would have to achieve some balance between, on the one side, the purposes of the scientists and engineers engaged in the research, and, on the other side, the corporate concerns of the institutions in which they were working. That the public and private non-profit funding agencies also will bring a further array of policy goals to bear – both directly and indirectly – on the determination of the ABCA’s operating principles is to be expected. In that context, it is only realistic to acknowledge that policy questions about intellectual property rights protections are likely to emerge among those that prove to be most problematic for the participants in research collaborations that receive substantial non-commercial (public sector and charitable foundation) support. This Part therefore offers some concluding observations on the issues raised in that connection, and the way in which they might best be resolved through the agency of the Advisory Board.

By emphasising the need to devise a new, flexible process for the ‘bottom up’ construction of institutional arrangements that will promote and support collaborative e-Science research, priority has already been accorded here to the public sector goal of rapidly and efficiently advancing scientific and technological knowledge. This position may be contrasted with according priority to the goal of capturing ‘private’ economic benefits from possession of new additions to the body of knowledge. The approach accordingly is to be favoured over efforts to codify existing institutional agreements for publicly funded research in standard form contractual agreements – especially those which would simply carry over into the academic institutional sphere intellectual property rights provisions modelled on the legal agreements governing commercial R&D partnerships, research joint ventures and similar consortium arrangements. The arguments for the latter position thus goes well beyond the point that standard form agreements may or may not strike the right balance between access and incentive in certain types

of scientific endeavour. A still much more serious problem lies in the present imbalance that has appeared within the intellectual property regime - between the extent and strength of the protection being accorded to holders of private (monopoly) rights, and societal protection of the public domain or a protected 'research commons' in scientific and technical data and information.

4.1 Preserving the Effectiveness of 'Open Science'

There is today a growing consensus among informed observers that the dominant trend of the past two decades towards broadening and strengthening of legal protections for intellectual property rights, and the privatising of the sources of scientific and technical data and information, has now gone too far. This assessment pertains to the situation existing among the handful of rich, economically advanced countries that do most of the world's organised science and engineering research. It is by no means confined to concerns that also have been expressed about the adverse impacts of the global IPR regime upon the developing economies' access to new scientific knowledge and knowledge-intensive goods and services.⁷¹

In Britain, the European Union and the US, several influential organisations have issued statements calling for a re-consideration of the place of intellectual property rights in contemporary science and technology. They point to a number of unintended, yet nonetheless undesirable impacts of current intellectual property policies upon the effective conduct publicly funded, academic research collaborations in science and engineering.⁷² Especially notable in this regard are the Royal

⁷¹ On the latter, however, see *Integrating Intellectual Property Rights and Development Policy*, The Report of the Commission on Intellectual Property Rights (2nd ed, 2002) esp Ch 1, 5–7 <http://www.iprcommission.org/papers/word/final_report/reportwordfinal.doc>.

⁷² The same theme is emerging more strongly in the recent writings of academic lawyers and economists in the US. See, for example, James Boyle (ed), 'The Public Domain' (Winter/Spring 2003) 66 (1 and 2) *Law and Contemporary Problems* (Special Issue of the Collected Papers from the Duke University Conference, held November 2001); Arti K Rais and Rebecca S Eisenberg, 'Bayh-Dole Reform and the Progress of Biomedicine' (2003) 66(1) *Law and Contemporary Problems* <<http://ssrn.com/abstract?id=348343>>; P A David, 'Can 'Open Science' be Protected from the Evolving Regime of Intellectual Property Protections' (Fall 2003) *Journal of Institutional and Theoretical Economics*, forthcoming (Discussion Paper 02–29, Stanford Institute for Economic Policy Research, Stanford University, (July 2003) <http://siepr.stanford.edu/papers/papersauth_D-H.html>).

Society's report entitled *Keeping Science Open: The Effects of Intellectual Property Policy on the Conduct of Science* (April 2003), and the Report of the EC Research Directorate General's European Research Area Expert Group on *Strategic Use and Adaptation of Intellectual Property Rights Systems in Information and Communications Technologies-based Research* (March 2003). These same concerns also permeate the US National Academy of Sciences recent publication: *The Role of the Public Domain in Scientific and Technical Data and Information: Proceedings of a Symposium* (September 2003).

None of these assessments are one-sided; all acknowledge that the protection of intellectual property rights can stimulate useful discoveries and inventions by protecting creative work and investments in costly research and development efforts. Further, they recognise that the prospective award of monopoly rights in the exploitation of new ideas can elicit the disclosure of discoveries that might otherwise be kept secret, and on that account may contribute to stimulating further advances in useful knowledge. Nevertheless, they concur in expressing serious concerns about the potential adverse impacts on the culture and practice of academic open science of the legal innovations, and the changes of institutional policy among the PROs in response to the emphasis that Western governments during the past two decades have placed upon near-term goals of 'wealth creation' through research. They deem it necessary to emphasise that 'high quality research is the gateway both to advances in knowledge and the wealth creation based on science'; that the competitive pursuit of patent rights creates incentives for secrecy that generally will be inimical to the rapid advancement of knowledge; that intellectual property rights are a basis for the imposition of costs, and the threat of costs which 'can hinder the free flow of ideas and information upon which science thrives.'⁷³

Consequently, the recent report of the Royal Society recommends, *inter alia*, the clarification and harmonising of the presently ambiguous exemptions from infringement of the patent laws permitted to scientific work under the headings 'private and non-commercial' and 'experimental' use.⁷⁴ The same document explicitly calls for reversal of

⁷³ The Royal Society, *Keeping Science Open: The Effects of Intellectual Property Policy on the Conduct of Science*, (Policy document 02/03 (April 2003)) v <<http://www.royalsoc.ac.uk>>.

⁷⁴ See, The Royal Society, *Keeping Science Open*, (April 2003) especially section 3.21 (patents); section 4.11–4.20 (copyright); section 5.5 (databases).

the recent introduction in UK law – specifically in the statutes implementing the *Database Directive (1996)* and the *Copyright Directive (2001)* – of narrow and ill-defined limitations on the ‘fair dealing’ exceptions provided for research; in a departure from traditional practice, these exceptions are confined to research that has ‘non-commercial purposes.’⁷⁵ Specifically, the Royal Society’s report notes that in the law enacting the *Database Directive* in the UK (on 1 January 1998), the fair dealing exception for research (and education) permits only extraction and not re-utilisation of the protected contents. Given the failure of the exiting statute to accommodate the needs of the scientific community in regard to digital databases – a ‘tool’ that has acquired increasing importance in numerous research contexts, the Royal Society’s report recommends that the laws be revised, so that even research that might be regarded as having some commercial value would be exempted from infringement of the database right and copyright.

Salutary as we believe these and related recommendations to be, their immediate practical force is mitigated by the fact that effecting significant legal reform is more often than not a complex, highly politicised and uncertain undertaking. This is likely to be true particularly in the area of intellectual property rights, where determined opposition must be expected from entrenched business firms whose strategies are predicated upon perpetuation of existing legal arrangements, and where the international repercussions and concerns for programmes of ‘harmonisation’ – not to mention the conflicts between the interests of the industrially advanced and the developing nations of the world – are likely to frustrate rapid progress.

The usefulness of the recommendations contained in those reports is further circumscribed by their own recognition that the effects of the intellectual property rights regime on the progress of particular fields of research, and on specific projects, may be quite different. Revisions of provisions in the intellectual property rights statutes cannot readily accommodate the effects of contextual variations without become inordinately complicated. Moreover, if differential rules are introduced that are perceived to create advantages for some research areas, or types of collaborations, this may induce efforts to reconfigure projects – or at

⁷⁵ In US legal parlance, such exceptions (to copyright law) are referred to under the heading ‘fair use’.

least to configure the outward appearance in order to exploit such advantages. Lastly, it is relevant to bear in mind that the foregoing proposals for legal reform would at best not address the entire problem; intellectual property law represents only one among the regulatory thickets on the institutional obstacle course and through which researchers attempting to advance a particular collaborative project need to find a feasible path.

4.2 Timely Action Through Contractually Constructed Collaboration Agreements

Consequently, it is our view that those who seek to advance collaborative e-Science in the here and now will be more effective if their attention and efforts focus, not on the codified details of intellectual property law, but upon the specific institutional structures that can be created contractually to facilitate particular collaborative structures that are most suitable to for the work of specific scientific projects. This approach, based upon the development of standard contractual clauses, is congruent with the conclusion of legal scholars in the US who advocate ‘contractually reconstructing the research commons’ as a way to mitigate the adverse effects of ‘a highly projectionist intellectual property environment.’⁷⁶ It also bears some kinship with recent initiatives to encourage the creative reuse of copyright protected material by providing a variety of readily implemented contractual alternatives to the full set of rights available to copyright owners under prevailing statutes.⁷⁷ In the present context, contractual construction of an e-Science research commons does not require that collaboration agreements must be created *de novo* for every occasion; the principles of modularisation and component

⁷⁶ See in particular, J H Reichman and P F Uhlir, ‘A Contractually Reconstructed Research Commons for Scientific Data in a Highly Protectionist Intellectual Property Environment’ (Winter/Spring 2003) 66 (1 and 2) *Law and Contemporary Problems* 315–462.

⁷⁷ A related contractual approach, utilising a menu of machine-readable copyright licenses, has been implemented by *Creative Commons*, a non-profit organisation developed on the initiative of Professor Lawrence Lessig (Stanford University, which describes itself as ‘dedicated to the creative reuse of copyrighted material’). For description of collaborative projects and organisations that are cooperating with Creative Commons, see <<http://creativecommons.org/learn/collaborators>>.

standardisation may be applied as usefully in this sphere as they are in the art of software engineering itself.

A concrete illustration may be given of the way that contractual clauses can be used in conjunction with the licensing terms of copyright protected software to solve difficult problems involving the balancing of different public policy interests. In the immediate context of the UK e-Science Programme it is germane to point out how contracts can work with existing law protecting intellectual property rights to accommodate seeming conflicting objectives of public funding agencies (and PROs), specifically in regard to the licensing of middleware and software applications tools that are being developed under the EPSRC e-Science Core Programme.

Much of the near-term rationale underlying investment in the development of that part of the technical ‘cyberinfrastructure’ rests on the contention that if access to data, information and facilities can be made easier, and less costly, there will be very substantial efficiency gains from the collaborative search for scientific knowledge. The economic case for reinforcing the ‘open science’ mode of collaboration is especially compelling in this area, given the modularity of well-engineered software, and the possibilities of generating recombinant novelty through re-use of already developed sub-routines. This purpose would be served by mandating the distribution of publicly funded code as free and open source software, making use of the terms of the already widely used GNU General Public License (GNU GPL).

On the other hand, both in government policy circles, among university administrators, and individual members of the academic research community, it is held to be highly desirable that the knowledge and information-goods generated by publicly funded research be available as a basis for private sector investment in its further elaboration and commercial distribution. To attain the latter goal by means of permitting PROs and their employees to exercise proprietary rights in software developed with substantial public support, however, would conflict with the rationale for public funding of this kind of R&D: the grant of copyright monopolies (and of software patents in some jurisdictions) if proved effective, would raise the economic costs of utilising the information-goods in question. This means it would raise their cost both as final goods, and as inputs for the production of further software innovations, including those by the private sector.

A pragmatic solution to this policy dilemma may be available. The proposal has been advanced to allow both goals to the goals to be served concurrently in some degree, by means of private contracts permitting modifications and further developments based upon publicly funded code released as ‘free and open source software.’⁷⁸ This ingenious use of features of the GNU Public License would call for public funding agencies in the first instance to mandate that all software created by their research projects be released under GNU General Public License (requiring distribution of the source codes if along with the machine code, among other terms of this standard license); and second, to allow the original copyright holder of such programs assign the copyright to some non-profit foundation or other entity that would oversee the granting of private contracts allowing modifications and elaboration upon the GPL’d code for ultimate commercial distribution.

The details and interpretation of the legal aspects of the GNU GPL that would permit this form of ‘dual licensing’ are interesting, but they need not be entered into here. This proposal has been injected into the present discussion primarily to illustrate the point that the development of contractual clauses (in this instance for commercial exploitation of the original extensions based upon GPL’d code) can be a useful device in the hands of collaborative public science communities, permitting collaborative projects between PRO’s and private sector firms. In addition, this concrete illustration serves to underscore the observation that the practical implementation of a novel proposal of this kind nonetheless would need to address many of the same institutional issues that have been examined by the preceding parts of this chapter. Consider just the following two sets of questions about the non-trivial practical details of implementing a program of copyleft and dual licensing for publicly funded software:

- a) Should a special foundation be created to hold the IP rights on publicly funded software, and should each funding agency have their own such foundation, or should they designate an independent national or international entity to which the original copyright author (and his/her host institution) would be

⁷⁸ See R A Ghosh, *Copyleft and dual licensing for publicly funded software development*, Draft version (1.0), MERIT/Institute of Infonomics, University of Maastricht (July 2003).

required ultimately to assign the ownership rights? Is an international or transnational foundation politically feasible? What about charities, such as the Wellcome Trust, or the Rockefeller Institutes - would they too have to form special foundations to fulfill this function, or would they be expected to voluntarily require assignment of rights to some pre-existing foundation(s)? How would such foundations be funded - by retaining a small proportion of the royalties garnered under from the private contracts that they issued? Would that create an institutional motivation to 'market' such permission on revenue-maximising terms, and if it did would the consequences be desirable?

- b) Are the foundation(s) to which publicly funded software would be assigned also to be made responsible for negotiating the equivalent of 'cross-licensing arrangements' affecting these private contractual permissions? What would be the mechanism for resolving negotiating conflicts among individual foundations attached to different funding agencies and countries? What would be the sources of such conflicts, and could those be suppressed by centralising at least the national assignment of GNU GPL licenses to a single entity? To the extent that the 'dual licenses' for private exploitation of GPL code are really contracts, how would one deal with issues of harmonisation among the jurisdictions in which such contract can be enforced?⁷⁹ How can those issues be prevented from obstructing contracting and cross-licensing agreements among those contract holders?

⁷⁹ This is not only a matter of differences among national legal jurisdictions, for, in the US, even though there is a uniform federal commercial code contract law is a matter for the State courts. Copyright, and hence 'copyleft' licenses avoid this problem because IPR is protected under Federal statutes.

Hence, it may be seen that one of the functions that the assignee foundations could perform would be that of establishing uniform contractual formulae, including jurisdiction-setting rules for the private contractual agreements. Alternatively, the ABCA or the pre-cursor Working Party (both of which were proposed in Part 3 of this chapter) could be tasked to provide suitable standard contractual clauses for this purpose. A further task of some importance would be to achieve some degree of harmonisation of the rules imposed by the funding agencies upon the initial holders of copyrights. Would one want the same search algorithm, or encryption programme to be differently 'regulated' in regard to its commercial exploitation, simply because it had been developed under a bioinformatics project supported by the Wellcome Trust, rather than by an EPSRC middleware development project? Probably not, but, as has already been noted, a policy of 'one size must fit all' would run the risk of removing flexibility and accommodation to the different realities of commercial exploitation opportunities, as well as the characteristics of different classes of software. This vexed issue of the appropriate degree of standardisation of the contractual rules for 'dual licensing' is one that certainly could be referred to the proposed Advisory Board, as one more problem to be addressed in the course of its work - along with others of a kindred nature.

From the foregoing it may be concluded that the services of an expert advisory body will be needed to deal with the many interrelated issues that arise in just this one area connection, even were it to be pre-determined that middleware developed by the e-Science Core Program's projects would all be released under the GNU General Public License. The precise form of the novel institutional that has been envisaged here as an Advisory Board on Collaboration Agreements is not what really matters, although features that assured its independence would be essential. What is required to meet the challenges of adaptive design of an appropriate institutional infrastructure, above all, is a guiding, architectural vision, and sufficient resources to mobilise and maintain the necessary technical expertise: first, to select and standardise the contractual components, and then to assess the performance of the various collaboration agreement that they have been used to construct. An entity able to sustain and assure continuity to those two, intertwining tasks ultimately could exert a powerful influence towards realising the

global promise of advancing knowledge and improving human welfare through e-Science.

5. CONCLUSION: THE CHALLENGE OF BUILDING NEW INSTITUTIONAL INFRASTRUCTURES FOR COLLABORATIVE e-SCIENCE

The most cursory review of modern sciences' dependence upon distributed digital data and information resources, and their growing needs for distributed and pervasive computing resources, suffices to reveal why so many distinct research communities view the success of technical efforts to provide an advanced 'digital infrastructure' as a common priority item on their respective requirements lists. To be sure, there are differences in the degrees of enthusiasm expressed about this goal and a number of valid questions that can be raised as to whether or not 'the Grid' is really of equally critical importance for the conduct of 21st century research in all the principal domain sciences, let alone mathematics or the social sciences. But that is only one, and surely not the most important, among the 'reality checks' that should be undertaken before committing extensive resources to the quest for Grid-enabled collaborative science as the lead user of the global cyberinfrastructure.

By comparison with the pace of engineering advances, far greater uncertainties continue to surround the extent to which individuals, groups, and organisations engaged in scientific and technical research are able to arrive at informal and formal contractual arrangements and institutionalised procedures to reduce the transactions costs of collaboration. The roots of this state of affairs lie in the micro and meso-level incentive structures formed by familiar features of the established legal and administrative regimes. Mundane as these obstacles may be, those transaction costs, and the economic rents protected by intellectual property rights that now occasion greater difficulties in negotiating agreements governing interorganisational research collaboration, cause private costs to greatly exceed the marginal social costs of effective access to data, information and information tools.

Economic analysis tells us that efficient resource allocation can occur in a decentralised regime when the prices of the goods in question are set equal to their marginal social costs. This implies that under modern

conditions, the imposition of substantial costs of access to existing data and information goods is tantamount to an inefficient tax, resulting in the wastage of society's resources. That burden is particularly difficult to justify on economic or ethical grounds where the initial, fixed costs of generating the information already has been borne by society through the provision of public funding for research and scholarship. Reducing the size of the transaction cost 'wedges', and the rents that are protected by intellectual property rights over scientific and technical data and information, is therefore an important challenge that must be met - if global research communities, and society more generally, are to benefit fully from the novel 'technologies of collaboration' that now are becoming engineering practicalities.

The same class of 'soft' problems underlie the exacting technical challenges that have emerged as serious obstacles to the commercial provision of Grid services *in interorganisational contexts*. Although the private incentives for overcoming those problems in the commercial sphere may be stronger than those felt by policy-makers with responsibilities for public sector research, the latter domain – for all its complexities – remains the more hospitable of the two environments for experimentation with new approaches to solving these problems. This is the case both because the ethos of cooperation in the collective pursuit of knowledge and the informal norms of 'open science' still persist in many research communities and because the public funding agencies still retain an important degree of policy-setting leverage over the relevant research organisations and institutions.

Consequently, it has been argued here that serious efforts should be made to explore some of the proposed modalities for the construction of an appropriate institutional infrastructure for collaborative e-Science. Not only may these yield direct benefits in terms of advancing the state of foundational scientific and engineering knowledge, but there can be significant spillovers. Experimentation with new institutional and organisation arrangements may yield solutions that find application to other fields of collaborative production that are both information intensive and regularly transcend the organisational boundaries.

Of course, it would be desirable for such governmental agencies and public research institutions to coordinate on policies that would promote 'bottom-up' initiatives for collaboration within the research communities, by more rationally managing publicly (and charitable,

quasi-public) funded data and information production and distribution in the rapidly progressing digitally networked research environment.⁸⁰ Recent proposals of this sort have been advanced for adoption by government agencies, featuring a variety of measures, including the following: (1) funding of public domain or open access data centres and active archives of foundational data sets derived from publicly supported research; (2) mandating open access to the scientific data and materials needed to replicate published results, and promoting open access to those results when they have issues from government-funded research projects; (3) providing for regular review and enforcement of research contract and grant clauses regarding open data availability, as an essential component of the public research infrastructure; and (4) protecting the interests of research users by developing open access principles and contractual provisions for licensing data products and services to or from the private sector, and for privatising the publication of essential government information.

But efforts to coordinate government policies along those lines are not sufficient. They can and should be conjoined with independent initiatives to address the immediate practical challenge of devising and adapting new institutional mechanisms that will reduce the myriad obstacles that add to the transactions costs and restrict the terms of interorganisational agreements within which collaborative research is hosted by public and charitable research organisations. Fortunately, there already are some encouraging movements in this direction. Independent foundations such as those emerging in the field of 'free and open source software' licensing and private initiatives, such as the Science Commons project recently launched by the non-profit corporation Creative Commons, are focused on providing research communities with licensing contracts formulated to facilitate the 'some

⁸⁰ For further elaboration, see P A David and P F Uhler, *Broadening the Information Commons for Science and Innovation: Strategic Institutional and Public Policy Approaches*, Proposal for the Planning Committee on the 2005 CODATA-ICSTI-U.S. NAS Workshop (May 18, 2004); P A David and P W Uhler, 'Creating the Information Commons for e-Science: An International Workshop Plan and Rationale' (July 2005) 91 *Codata Newsletter* <<http://www.codataweb.org/UNESCOmtg/workshopplan.html>>.

rights reserved' sharing of scientific information, data, and research materials.⁸¹

The negotiation of agreements that can clear researchers' paths through 'patent thickets,' 'database barricades,' and 'copyright stacks' is a vital task, but it is only one part of the necessary work— as the preceding pages have sought to show. The complexities and uncertainties of modern scientific research, and the multiplicity of the participating agents and agencies that global e-Science will involve, call for a more comprehensive 'bottom-up' approach to the contractual reconstruction and expansion of the scientific commons. The proposed development of suites of modular contractual clauses, and guidelines for informal cooperative procedures that would enable construction of a variety of customised, flexible 'collaboration agreements', appear to offer a practical 'way forward' for public funding agencies to encourage and endorse.

In closing, as bromidic and predictable as the academic's closing plea for 'further research' may be, surely it will be accepted as warranted in the present connection. There is a largely unmet need for empirical assessments of the nature and severity of the varied impediments to an effectively functioning infrastructure for publicly supported scientific and technological collaborations in specific research domains. Intrinsically interesting methodological challenges as well as difficult data collection tasks lie along the route to systematic measures of the effects of the incentives and constraints of such undertakings that are created by prevailing organisational norms, institutional rules, and governmental policies. A better understanding of their differential impacts upon the direction and conduct of research projects in the various domain sciences and upon exploratory work in emerging transdisciplinary fields would be of real value in identifying specific targets for remedial attention. Only on the basis of such knowledge will it be practical to formulate and implement coordinated strategies of private and public

⁸¹ Efforts of this kind are very much in line with the pragmatic spirit of Reichman and Uhler's (2003) advocacy of efforts to 'contractually reconstruct the science commons' in an environment characterised by increasingly strong and pervasive intellectual property rights protections. More specific details about the programs being undertaken by Science Commons and its relationship to Creative Commons is available at <<http://sciencecommons.org>>. It is appropriate here to disclose the 'interest' of one of the authors — David is a member of the Scientific Advisory Board for Science Commons.

action that have a good prospect of freeing distributed collaborative research from the persisting constraints of the present mal-adapted institutional infrastructure.

THE LAW AS CYBER INFRASTRUCTURE

Professor Brian Fitzgerald and Kylie Pappalardo¹

In almost everything we do, the law is present. However, we know that strict adherence to the law is not always observed for a variety of pragmatic reasons. Nevertheless, we also understand that we ignore the law at our own risk and sometimes we will suffer a consequence.

In the realm of collaborative endeavour through networked cyberinfrastructure we know the law is not too far away. But we also know that a paranoid obsession with it will cause inefficiency and stifle the true spirit of research. The key for the lawyers is to understand and implement a legal framework that can work with the power of the technology to disseminate knowledge in such a way that it does not seem a barrier. This is difficult in any universal sense but not totally impossible. In this article, we will show how the law is responding as a positive agent to facilitate the sharing of knowledge in the cyberinfrastructure world.

One general approach is to develop legal tools that can provide a generic permission or clearance of legal rights (for example, copyright or patent) in advance (usually subject to conditions) that can be implemented before or at the point of use. This has become known as open licensing and will be discussed below in terms of copyright and patented subject matter.²

¹ Professor of Intellectual Property and Innovation (QUT) and Research Officer OAK Law Project, QUT Law Faculty (respectively). This chapter was first published as an article in (2007) 3(3) *CTWatch Quarterly* 61–7 <<http://www.ctwatch.org/quarterly/pdf/ctwatchquarterly-12.pdf>>.

² For more information, see Professor Brian Fitzgerald et al, *OAK Law Project Report No 1: Creating a Legal Framework for Copyright Management of Open Access Within the Australian Academic and Research Sector* (2006) <http://eprints.qut.edu.au/archive/00007306/01/Printed_Oak_Law_Project_Report_No_1.pdf>; and Dr Anne Fitzgerald and Kylie Pappalardo, *Building the Infrastructure for Data Access and Reuse in Collaborative Research: An Analysis of the Legal Context* (2007) <http://www.oaklaw.qut.edu.au/files/Data_Report_final_web.pdf>.

However, open licensing will not be adopted by everyone nor in every situation is it suitable. A generalisation is that it will be advocated in the context of publicly funded research producing tools and knowledge upon which platform technologies are built where considerations such as privacy are not an issue.

Where open licensing is not being used, the many parties to a collaborative endeavour will normally be required to map the scope and risk of their mutual endeavour through a contract. Contracts can take time to negotiate and, in many instances, promise to frustrate the fast paced and serendipitous nature of research fuelled by high powered cyberinfrastructure. To this end a number of projects throughout the world, for example The Lambert Project in the UK,³ the University Industry Demonstration Project (UIDP) in the USA,⁴ and (amongst other projects) the 7th Framework Project in the EU,⁵ have begun asking how we might be able to improve this situation. Suggestions include standard form or off the shelf contracts covering a variety of situations, a database of key clauses and, in the case of the UIDP project, a software based negotiation tool called the Turbo-Negotiator. Legal instruments that can match the dynamic of the technology and appear seamless and non-invasive are the goal. More work in this area is needed (and happening) and is critical to ensuring we have the law and technology of cyberinfrastructure working to complement each other.

In the remainder of this article we will focus on the open licensing model.

OPEN LICENSING

Open Content Licensing

From a legal perspective, one of the most significant responses to the technological advances that have revolutionised the creation and

³ Lambert Working Group on Intellectual Property
<<http://www.innovation.gov.uk/lambertagreements/index.asp?lv1=1&lv2=0&lv3=0&lv4=0>>.

⁴ University Industry Demonstration Partnership <<http://uidp.org/>>.

⁵ Seventh Research Framework Programme (FP7)
<http://cordis.europa.eu/fp7/home_en.html>. See further, European Commission, *Intellectual Property and Technology Transfer* <http://ec.europa.eu/invest-in-research/policy/ipr_en.htm>.

distribution of copyright materials during the last decade has been the development of new systems for licensing (or authorising) others to obtain access to and make use of the protected material. These new forms of licences – usually referred to as ‘open content’ – are founded upon an acknowledgement of the existence of copyright in materials embodying knowledge and information, but differ from licences commonly used before the advent of the digital era in key respects. As well as being relatively short, simple and easy to read, they are standardised, conceptually interoperable with other open content licences, machine (computer) enabled and have the advantage that, since they are automated and do not require negotiation, they eliminate (or at least minimise) transaction costs. Running with the copyright material to which they are attached (thereby avoiding the privity issue where rights are conferred contractually), open content licences identify materials that are available for reuse and grant permissive rights to users, thereby facilitating access and dissemination.⁶ The most widely used of the open content licences are the Creative Commons licences.⁷ These licences attach to the copyright material and provide that anyone can reuse the material subject to giving attribution to the author of the material and subject to any of the optional conditions as selected by the licensor. The optional conditions are:

- non-commercial use;
- no derivative materials based on the licensed material are to be made; or
- share alike – others may distribute derivative materials based on the licensed material, but only under a licence identical to that covering the licensed material.

Creative Commons licences have more commonly been applied to publications than to research data. They have been particularly useful for academic authors depositing their publications in university or scholarly digital repositories or databases. Repositories help to make publications more accessible to the research and general communities.

⁶ B Fitzgerald et al, *OAK Law Project Report No 1: Creating a Legal Framework for Copyright Management of Open Access Within the Australian Academic and Research Sector* (2006) [1.22] <http://eprints.qut.edu.au/archive/00007306/01/Printed_Oak_Law_Project_Report_No_1.pdf>.

⁷ Creative Commons <<http://creativecommons.org/>>.

The advantage of a Creative Commons licence is that it tells people accessing the publication what they can and cannot do with the material, without the copyright owner having to deal with permissions on a case-by-case basis.

Below are two examples of scientific research publication projects that promote open access and reuse of material by utilising open content licensing models.

Example One – PLoS ONE

The Public Library of Science (PLoS) is a non-profit, open access, scientific publishing project that aims to create a library of peer-reviewed scientific and medical journals that are made available online without restrictions under open content licences.⁸ PLoS ONE is a peer-reviewed, scientific literature journal that enables scientific research to be published and disseminated within weeks, avoiding delays associated with traditional means of publication.⁹

The features of PLoS ONE include:

- *rapid publication* – realising that the rapid publication and dissemination of research is one of the highest priorities, PLoS ONE ensures a streamlined electronic production workflow that ensures papers are published within weeks of submission;
- *freedom of use and ownership* – in accordance with the CC attribution licence, PLoS ONE enables users to read, copy, distribute and share papers freely without restrictions and formal permission, provided that the original author and source are cited; and
- *high impact* – PLoS ONE has been designed in light of the fact that papers published in OA journals are more likely to be read and cited given the lack of barriers to access.

Example Two – Nature Precedings

Nature Precedings is an online database designed to allow scientific researchers to share pre-publication research, unpublished manuscripts, presentations, white papers, technical papers, supplementary findings

⁸ Public Library of Science (PLoS) <<http://www.plos.org>>.

⁹ PLoS One <<http://www.plosone.org>>.

and other scientific documents.¹⁰ Contributions are taken from biology, medicine (except clinical trials), chemistry and earth sciences. The database is free of charge to access and use, and is intended to provide a rapid means of disseminating emerging results and new theories, soliciting opinions and recording the provenance of ideas.

Nature Precedings aims to make scientific documents citable, globally available and stably archived. To this end, it can also be used as an archiving tool for scientists to store their work for their own future convenience.

Submissions made to Nature Precedings are screened by a professional curation team for relevance and quality, but are not subject to peer review. The database is designed to complement scientific journals by providing a more rapid and informal communication system, but submissions to Nature Precedings are not subject to the same rigorous and time-consuming reviews as submissions made to scientific journals.

The Nature Precedings website states that scientists should own copyright in a document and have permission from other copyright holders (for example, co-authors), before they submit the document to Nature Precedings.¹¹ Copyright then remains with the author. However, the website encourages scientists to release their work under a Creative Commons Attribution Licence so that content can be quoted, copied and disseminated, provided that the original source is correctly cited.¹²

Authors who own copyright in their publication will be able to place a Creative Commons licence on their work, but if they have assigned copyright to their publisher or another party, they will need to ask permission from that party before they can attach a Creative Commons licence. A problem that often arises in this situation is that authors are unsure of whether they own copyright or their publisher owns copyright. Even when authors know that they have transferred copyright to their publisher, they may be reluctant to ask their publisher if they can attach

¹⁰ Nature Precedings <<http://precedings.nature.com/>>.

¹¹ Nature Precedings, *Copyright* <<http://precedings.nature.com/about>>.

¹² Nature Precedings, *Copyright* <<http://precedings.nature.com/about>>.

a Creative Commons licence to their work for fear of jeopardising their relationship with the publisher.¹³

These issues are best dealt with through established policies. Every research and academic institution should have in place policies relating to copyright management, including the licensing of copyright works. These policies should deal with the legal impediments to making copyright material openly accessible, including determining who owns copyright, how to obtain necessary permissions from copyright owners and how to licence material in a way that grants the appropriate rights but retains the appropriate controls. The policies may also deal with non-legal issues, including how to get authors interested in open access repositories and how to assist authors in maintaining a positive relationship with their publisher while asserting additional rights.¹⁴

The Creative Commons open content principles have been extended to the sharing of scientific data and publications through the Science Commons Project.¹⁵ As explained on the Science Commons website, Creative Commons licences can be used in relation to databases that attract copyright protection.¹⁶ An example of a database that uses a Creative Commons licence appears below.

Example Three – UniProtKB/Swiss-Prot Protein Knowledgebase

UniProtKB/Swiss-Prot is a protein knowledgebase established in 1986 and maintained since 2003 by the UniProt Consortium. The UniProt Consortium is a collaboration between the Swiss Institute of Bioinformatics and the Department of Bioinformatics and Structural Biology of the Geneva University, the European Bioinformatics Institute (EBI) and the Georgetown University Medical Centre's Protein Information Resource.

The data held within UniProtKB includes protein sequences, current knowledge on each protein, core data (sequence data; bibliographical references and taxonomic data) and further annotation. The database is

¹³ For more information, see Kylie Pappalardo and Dr Anne Fitzgerald, *A Guide to Developing Open Access Through Your Digital Repository* (2007) <<http://www.oaklaw.qut.edu.au/node/32>>.

¹⁴ For more information, see Kylie Pappalardo and Dr Anne Fitzgerald, *A Guide to Developing Open Access Through Your Digital Repository* (2007) <<http://www.oaklaw.qut.edu.au/node/32>>.

¹⁵ Science Commons <<http://sciencecommons.org/>>.

¹⁶ See <<http://sciencecommons.org/resources/faq/databases>>.

organised through a web interface that displays the data associated with each protein sequence.

The UniProt Consortium states that the public databases maintained by UniProt Consortium members are freely available to any individual and for any purpose.

A copyright statement on the UniProtKB website states:

We have chosen to apply the Creative Commons Attribution-NoDerivs Licence to all copyrightable parts of our databases. This means that you are free to copy, distribute, display and make commercial use of these databases, provided you give us credit. However, if you intend to distribute a modified version of one of our databases, you must ask us for permission first.¹⁷

The UniProtKB open access system has been described as operating on an 'honour system' on the basis that the user community is small and so accurately monitored by electronic tracking that non-compliance with the copyright licence would risk unacceptable costs in loss of reputation, peer pressure and possible denial of privileges.

Open Patent Licensing

Increased interest in sharing data also raises issues in relation to patents. Patents protect products and processes that are novel, useful and involve an inventive or innovative step. Patents must be registered and confer on the patentee the exclusive right to use or sell the patented product during a certain period of time (usually 20 years).

For researchers intending to seek patent protection for inventions derived from their research, a primary concern is whether they will be able to obtain a patent and whether disclosure of their data to other researchers could prevent them from obtaining a patent (because the product would no longer be 'novel'). For researchers who do not intend to patent, a concern is whether another person could secure a patent over an invention that encompasses the researcher's data.

Some researchers will be more interested in making their data openly available to advance research than in commercialising patented products

¹⁷ UniProtKB <<http://www.uniprot.org/terms>>.

or processes derived from their research. These researchers will not be concerned that public disclosure of their research data could prevent them from obtaining a patent because the invention is no longer novel or is obvious. However, disclosure of data, in itself, will not always be enough to prevent patenting. The problem arising from the public release of data is that it leaves the way open for another party to make improvements to the disclosed data and then make those improvements proprietary.

Claire Driscoll of the NIH describes the dilemma as follows:

It would be theoretically possible for an unscrupulous company or entity to add on a trivial amount of information to the published ... data and then attempt to secure 'parasitic' patent claims such that all others would be prohibited from using the original public data.¹⁸

Where information or data is used to develop a patentable invention, the subsequent patent rights may be broad enough to cover use of the actual data forming part of the invention. As Eisenberg and Rai explain:

Although raw genomic data would not undermine claims to specific genes of identified function, annotated data might do so. A major goal of annotation is to identify coding regions in the genome and add information about the function of the protein for which the region codes.¹⁹

Consequently, some research projects have relied on licensing methods, similar to the open content copyright licences described above, in an attempt to keep the data 'open', rather than simply releasing the data into the public domain.

One example is the HapMap Project, which required anyone seeking to use research data in the HapMap database to first register online and enter into a click-wrap licence for use of the data. The licence prohibited licensees from filing patent applications that contained claims

¹⁸ C Driscoll, 'NIH data and resource sharing, data release and intellectual property policies for genomics community resource projects' (2005) 15(1) *Expert Opinion on Therapeutic Patents* 4.

¹⁹ R Eisenberg and A Rai, 'Harnessing and Sharing the Benefits of State-Sponsored Research: IP Rights and Data Sharing in California's Stem Cell Initiative' (2006) 21 *Berkeley Law Journal* 1187, 1202.

to particular uses of data obtained from the HapMap databases, unless that claim did not restrict the ability of others to freely use the data.²⁰

Another approach – currently being practised by the CAMBIA project – is to obtain a patent and then open licence the use of the patented invention on certain conditions. Some argue that, in specific areas, effective open access will only be achieved by allowing a certain level of use of the copyright and patented material.

The CAMBIA Approach

CAMBIA is an international, independent, non-profit research institute led by well known scientist, Richard Jefferson. CAMBIA was designed to ‘foster innovation and a spirit of collaboration in the life sciences’.²¹ This goal is achieved through four interconnected work products:

- *Patent Lens*, which provides tools to make patents and patent landscapes more transparent;
- *Biological Open Source Initiative (BiOS)*, which advocates for the sharing of life sciences technology and data through a series of licences;
- *BioForge*, a research portal (or repository) that makes data and technologies openly available for others to use in new innovations, whether for research, commercial use, or humanitarian use; and
- *CAMBIA’s Materials*, new technologies developed by CAMBIA, particularly in the field of genetics, which CAMBIA makes openly available under a BiOS licence.

CAMBIA has also applied for and obtained 12 patents of biological material in different patent offices around the world. CAMBIA’s approach involves obtaining patents over products or processes, but then licensing the use of those inventions under open terms. Another object is to encourage innovation. CAMBIA

²⁰ HapMap, *Project Public Access Licence*, previously at <<http://www.hapmap.org/cgi-perl/registration>>. Users are no longer required to enter into this licence to use the HapMap database. See also R Eisenberg and A Rai, ‘Harnessing and Sharing the Benefits of State-Sponsored Research: IP Rights and Data Sharing in California’s Stem Cell Initiative’ (2006) 21 *Berkeley Law Journal* 1187, 1202.

²¹ CAMBIA <<http://www.cambia.org/daisy/cambia/home.html>>.

Strives to create new norms and practices for dynamically designing and creating the tools of biological innovation, with binding covenants to protect and preserve their usefulness, while allowing diverse business models for wealth creation, using these tools.²²

CAMBIA has developed two open licences relevant to data – the BiOS Plant Enabling Technology Licence and the BiOS Genetic Resource Technology Licence. Paragraph 2.1 of each licence gives licensees

A worldwide, non-exclusive, royalty-free right and licence to make and use the IP & Technology for the purpose of developing, making, using, and commercializing BiOS Licensed Products without obligation to CAMBIA, including a sub-licence ...²³

This gives licensees the right to sub-licence the material, as long as it is sub-licensed under the same terms as contained in the original licence agreement.

CAMBIA's model allows researchers to obtain patents over inventions that build upon CAMBIA's research data. However, instead of using patent licences to 'extract a financial return from a user of a technology', CAMBIA advocates using a patent licence to 'impose a covenant of behaviour'.²⁴

According to CAMBIA, the purpose of the BiOS licences is that:

Instead of royalties, BiOS licensees must agree to legally binding conditions in order to obtain a licence and access to the protected commons. These conditions are that improvements are shared and that licensees cannot appropriate the fundamental 'kernel' of the technology and

²² R Jefferson, 'Science as Social Enterprise: The CAMBIA BiOS Initiative' (2006) 13 *Innovations* 22
<http://www.bios.net/daisy/bios/3067/version/default/part/AttachmentData/data/INNOV0104_pp13-44_innovations-in-practice_jefferson.pdf>.

²³ CAMBIA, Biological Innovation for Open Science, *About BiOS (Biological Open Source) Licences*, <<http://www.bios.net/daisy/bios/398>>.

²⁴ R Jefferson, 'Science as Social Enterprise: The CAMBIA BiOS Initiative' (2006) 13 *Innovations* 22
<http://www.bios.net/daisy/bios/3067/version/default/part/AttachmentData/data/INNOV0104_pp13-44_innovations-in-practice_jefferson.pdf>.

improvements exclusively for themselves. Licensees obtain access to improvements and other information, such as regulatory and biosafety data, shared by other licensees. To maintain legal access to the technology, licensees must agree not to prevent other licensees from using the technology in the development of different products.²⁵

By making the licence cost-free, CAMBIA hopes to encourage what founder Richard Jefferson terms '[t]he most valuable contribution to the license community: 'freedom to innovate'''.²⁶

CAMBIA is currently developing a new version of the BiOS licence, which to our understanding will remove any positive obligation to share improvements in return for some type of covenant not to enforce rights in relation to patented improvements against members of the CAMBIA community.

CONCLUSION

Any research project should adopt a 'mission-driven approach'. The question to be asked is, 'what do we want to achieve?' The goal may be commercial gain, or it may be the advancement of research for the public good, or both. Open access to research data and publications should always be considered, especially in the case of publicly funded research.²⁷ The level of access to and reuse of research data and publications that is to be allowed should ideally be determined at the outset of a research project.

From the commencement of a research project, it is imperative to have appropriate policies and frameworks in place. Policies must cover copyright management and data management. Copyright management policies should deal with copyright ownership rights and how copyright

²⁵ R Jefferson, 'Science as Social Enterprise: The CAMBIA BiOS Initiative' (2006) 13 *Innovations* 22

<http://www.bios.net/daisy/bios/3067/version/default/part/AttachmentData/data/INNOV0104_pp13-44_innovations-in-practice_jefferson.pdf>.

²⁶ R Jefferson, 'Science as Social Enterprise: The CAMBIA BiOS Initiative' (2006) 13 *Innovations* 22

<http://www.bios.net/daisy/bios/3067/version/default/part/AttachmentData/data/INNOV0104_pp13-44_innovations-in-practice_jefferson.pdf>.

²⁷ OECD, *Declaration on Access to Research Data from Public Funding*

<http://www.oecd.org/document/0,2340,en_2649_34487_25998799_1_1_1_1,00.html>

protected material is to be shared. Researchers should consider the various open content licensing models that can be applied to their copyright material. Data management plans should deal with how data is to be generated, managed and stored; data ownership rights and legal controls that may apply to data (including patents); and how access will be provided to the data and how the data will be disseminated.

Interestingly, some argue that, while open access in terms of copyright material will allow us to read that material and potentially to reproduce and electronically communicate it to colleagues, it most likely will not provide permission to use or exploit related patented material. One of the challenges for the near future will be to consider to what extent open access to publicly funded knowledge (for example, that makes up tools or platform technologies in biotechnology) requires an accompanying commitment to allow a certain level of use of patented material. In this regard, the CAMBIA project provides an interesting approach that deserves close attention in coming years.

As lawyers, we hope that the law can adapt to facilitate the very great potential cyberinfrastructure promises us. To this end, we need to think of legal tools as being part of the infrastructure and work towards providing innovative models for the future.

PART THREE

DATA OWNERSHIP, ACCESS
AND REUSE

UNDERSTANDING THE LEGAL IMPLICATIONS OF DATA SHARING, ACCESS AND REUSE IN THE AUSTRALIAN RESEARCH LANDSCAPE¹

**Professor Anne Fitzgerald,² Kylie Pappalardo,³
Anthony Austin⁴**

INTRODUCTION

In today's world, researchers are increasingly involved in data-intensive research projects that cut across geographic and disciplinary borders.⁵ Quality research now often involves virtual communities of researchers participating in large-scale web-based collaborations, opening their early-stage research to the research community in order to encourage broader

¹ This chapter is derived from the publication: Dr Anne Fitzgerald and Kylie Pappalardo (with the assistance of Professor Brian Fitzgerald, Anthony Austin and others), *Building the Infrastructure for Data Access and Reuse in Collaborative Research: An Analysis of the Legal Context* (2007) OAK Law Project <<http://www.oaklaw.qut.edu.au/reports>>.

² Professor, Faculty of Law, Queensland University of Technology, JSD (*Columbia*), LLM (*Columbia*), LLM (*London*), LLB (Hons) (*Tasmania*), Grad Dip Welfare Law (*Tasmania*), BSW (*UQ*), Barrister (*Qld*), Practitioner – High Court of Australia, am.fitzgerald@qut.edu.au

³ Research Officer, Open Access to Knowledge (OAK) Law Project, Faculty of Law, Queensland University of Technology LLB (*QUT*) (Hons), BCI (Creative Writing) (*QUT*), Grad Dip Legal Prac (*QUT*), k.pappalardo@qut.edu.au

⁴ Research Officer, Open Access to Knowledge (OAK) Law Project, Faculty of Law, Queensland University of Technology, LLB (*QUT*), Grad Dip Legal Prac (*QUT*), LLM (*QUT*), Solicitor of the Supreme Court of Queensland and the Federal Court of Australia, anthony.austin@qut.edu.au

⁵ International Council for Science (ICSU), *Scientific Data and Information: A report of the CSPIR Assessment Panel* (2004) 7; see also Dr Anne Fitzgerald and Kylie Pappalardo, *Building the Infrastructure for Data Access and Reuse in Collaborative Research: An Analysis of the Legal Context* (2007) 6, OAK Law Project <<http://www.oaklaw.qut.edu.au/reports>> (hereinafter A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007)).

participation and accelerate discoveries.⁶ The result of such large-scale collaborations has been the production of ever-increasing amounts of data. In short, we are in the midst of a data deluge.⁷

Accompanying these developments has been a growing recognition that if the benefits of enhanced access to research are to be realised, it will be necessary to develop the systems and services that enable data to be managed and secured.⁸ It has also become apparent that to achieve seamless access to data it is necessary not only to adopt appropriate technical standards, practices and architecture, but also to develop legal frameworks that facilitate access to and use of research data.⁹

This chapter provides an overview of the current research landscape in Australia as it relates to the collection, management and sharing of research data. The chapter then explains the Australian legal regimes relevant to data, including copyright, patent, privacy, confidentiality and contract law. Finally, this chapter proposes the infrastructure elements that are required for the proper management of legal interests, ownership rights and rights to access and use data collected or generated by research projects.

THE AUSTRALIAN DATA LANDSCAPE

The last few years have seen a revolution in the way that research data is produced, stored, analysed and disseminated.¹⁰ Now, vast amounts of data can be generated and accessed through distributed networks online. In response to the enormous growth in data collection and generation in recent years, there has been increased interest from Australian government and research sectors in developing systems to manage data

⁶ A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 6.

⁷ In an interview with Richard Poynder, Tony Hey said, 'We are going to be deluged with data in almost every field': Richard Poynder, Interview with Tony Hey 'A Conversation with Microsoft's Tony Hey' *Open and Shut?* (Blog, 12 December 2006) <<http://poynder.blogspot.com/2006/12/conversation-with-microsofts-tony-hey.html>> at 5 May 2008.

⁸ A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 6–7.

⁹ A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 9.

¹⁰ A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 3.

and facilitate access to research outputs.¹¹ This section provides a brief overview of some of these initiatives.

Government Initiatives

In May 2004, then Prime Minister John Howard announced that the Australian Government would establish quality and accessibility frameworks for publicly funded research as part of the *Backing Australia's Ability – Building Our Future through Science and Innovation* package.¹² The Accessibility Framework for Publicly Funded Research was designed to manage research information, outputs and infrastructure in order to enable them to be more readily discovered, accessed and shared. It aims to provide a regulatory environment that both enables and encourages the population of digital repositories in order to provide better access to information.¹³

A project funded under the *Backing Australia's Ability* package is the National Collaborative Research Infrastructure Strategy (NCRIS). The NCRIS capability known as Platforms for Collaboration supports technological platforms that enhance researchers' ability to generate, collect, share, analyse, store and retrieve information.¹⁴ A central component of Platforms for Collaboration is the Australian e-Research Infrastructure Council (AeRIC), established by the federal Government's Department of Education Science and Training (DEST)¹⁵ upon the

¹¹ See for example, National Collaborative Research Infrastructure Strategy (NCRIS) Committee, Submission to the Review of the National Innovation System (NIS) (April 2008) and AeRIC, *Closing the Gap: Connecting Researchers to the Innovation System Through Sustained Investments in Collaborative Research Infrastructure*, Submission to the Review of the National Innovation System (NIS) (April 2008).

¹² See <http://www.dest.gov.au/sectors/research_sector/policies_issues_reviews/key_issues/accessibility_framework/> and <<http://backingaus.innovation.gov.au/>> at 24 April 2008.

¹³ See <http://www.dest.gov.au/sectors/research_sector/policies_issues_reviews/key_issues/accessibility_framework/> at 24 April 2008. See also A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 3.

¹⁴ <http://www.ncris.dest.gov.au/capabilities/collaborative_investment_plan_platforms.htm> at 24 April 2008; see also, National Collaborative Research Infrastructure Strategy (NCRIS) Committee, Submission to the Review of the National Innovation System (NIS) (April 2008).

¹⁵ Since the change of Federal Government, AeRIC now falls under the auspices of the Department of Innovation, Industry, Science and Research (DIISR): see Dr Rhys Francis on

recommendation of the Australian Government e-Research Coordinating Committee.¹⁶ In the report, *An Australian e-Research Strategy and Implementation Framework*, the e-Research Coordination Committee had recommended that:

the Government convenes a working group to develop an Australian Research Data Strategy that will support a standardised national approach to the management of data collected, generated and used by the Australian research community.¹⁷

This recommendation was endorsed in the NCRIS Platforms for Collaboration Final Investment Plan.¹⁸

AeRIC's responsibilities were established at its inaugural meeting on 23 July 2007, as:

ensur[ing] that world class services and expertise are identified, developed and delivered nationwide in ways that support effective e-Research within and across all research disciplines ... includ[ing] services and expertise relating to: data capture, management, retention, publication, discovery and reuse ...¹⁹

AeRIC undertakes an important coordination role in relation to the NCRIS Platforms for Collaboration infrastructure.²⁰ It is tasked with ensuring the integration and sustainability of research infrastructure and

behalf of Professor Tom Cochrane, AeRIC submission to the National Innovation System (NIS) Review – coversheet, 30 April 2008.

¹⁶ In October 2004, the Australian Government committed to the formation of an overarching e-Research Coordinating Committee, which would provide expert advice to the Government on a strategic framework for the development of Australia's e-Research capacity: see for example, Catherine Harboe-Ree, 'eResearch Coordinating Committee' (CAUL Presentation, September 2005) <<http://www.caul.edu.au/caul-doc/caul20052ereseach.ppt>> at 3 May 2008.

¹⁷ Final Report of the e-Research Coordinating Committee, *An Australian e-Research Strategy and Implementation Framework*, DEST, (April 2006) 55.

¹⁸ See <www.ncris.dest.gov.au/capabilities/documents/PfC_Investment_Plan_Summary_pdf.htm> at 24 April 2008. The NCRIS Committee accepted the Final Investment Plan on 13 April 2007: <<http://www.pfc.org.au/bin/view/Main/PlatformsHistory>> at 24 April 2008.

¹⁹ See <<http://www.pfc.org.au/bin/view/Main/AeRIC>> at 24 April 2008.

²⁰ AeRIC, *Closing the Gap: Connecting Researchers to the Innovation System Through Sustained Investments in Collaborative Research Infrastructure*, Submission to the Review of the National Innovation System (NIS) (April 2008) 9.

services capitalising on the Government's substantial investments in NCRIS capabilities.²¹

In October 2007, DEST and AeRIC released the report, *Towards the Australian Data Commons*,²² proposing the establishment of the Australian National Data Service (ANDS). A similar proposal had previously been put forward in the Platforms for Collaboration Final Investment Plan as a means of addressing the recommendations of the Prime Minister's Science Engineering and Innovation Council (PMSEIC) Data for Science Working Group²³ in its December 2006 report:

Recommendation 1: That Australia's government, science, research and business communities establish a nationally supported long-term strategic framework for scientific data management, including guiding principles, policies, best practices and infrastructure.

Recommendation 6: That the principle of open equitable access to publicly-funded scientific data be adopted wherever possible and that this principle be taken into consideration in the development of data for science policy and programmes. As part of this strategy, and to enable current and future data and information resources to be shared, mechanisms to enable the discovery of, and access to, data and information resources must be encouraged.

Recommendation 9: That in the context of developing the strategic framework for scientific data management, Australia's intellectual property approaches be checked to ensure they do not impede the sharing of data ...²⁴

²¹ AeRIC, *Closing the Gap: Connecting Researchers to the Innovation System Through Sustained Investments in Collaborative Research Infrastructure*, Submission to the Review of the National Innovation System (NIS) (April 2008) 9.

²² AeRIC, *Towards the Australian Data Commons: A proposal for an Australian National Data Service*, DEST (October 2007).

²³ See <www.ncris.dest.gov.au/capabilities/documents/PfC_Investment_Plan_Summary_pdf.htm> at 24 April 2008. The NCRIS Committee accepted the Final Investment Plan on 13 April 2007: <<http://www.pfc.org.au/bin/view/Main/PlatformsHistory>> at 24 April 2008.

²⁴ PMSEIC Working Group on Data for Science, *From Data to Wisdom: Pathways to Successful Data Management for Australian Science* (December 2006) 11–12.

ANDS offers common services in support of research data collections and integration infrastructure to facilitate sharing and reuse of data.²⁵ The ANDS Utility Program will provide a national registry covering issues such as data access policies, usage rights and licensing requirements associated with data access.²⁶ It will also provide template data access policies that can be adapted for discipline specific needs.²⁷

At an AeRIC meeting on 22 February 2008, it was reported that a contract was signed in November 2007 with Monash University to conduct the ANDS Establishment Project through to the end of June 2008.²⁸ Under this agreement, Monash University will work with the Australian National University (ANU), the Commonwealth Scientific and Industrial Research Organisation (CSIRO) and other relevant parties to develop the necessary elements to move to full ANDS implementation from July 2008.²⁹

Research projects

In addition to the larger scale initiatives described above, there is an abundance of smaller projects that focus on the collecting and compiling of research data in a specific scientific field. One example is the Integrated Marine Observing System (IMOS), which is coordinated by staff at the University of Tasmania supported by CSIRO Marine and Atmospheric Research.³⁰ IMOS is a nation-wide collaborative program designed to observe the oceans around Australia, including the coastal oceans and the 'bluewater' open oceans.³¹ One of the five IMOS research 'nodes' is the Great Barrier Reef Ocean Observing System (GBROOS), which is an observation network covering the eastern Coral

²⁵ AeRIC, *Towards the Australian Data Commons: A proposal for an Australian National Data Service*, DEST (October 2007) 4.

²⁶ AeRIC, *Towards the Australian Data Commons: A proposal for an Australian National Data Service*, DEST (October 2007) 36.

²⁷ AeRIC, *Towards the Australian Data Commons: A proposal for an Australian National Data Service*, DEST (October 2007) 36.

²⁸ AeRIC Executive Director's Report, *Meeting #5* (22 February 2008) <<http://www.pfc.org.au/bin/view/Main/AeRIC-5>> at 24 April 2008.

²⁹ AeRIC Executive Director's Report, *Meeting #5* (22 February 2008) <<http://www.pfc.org.au/bin/view/Main/AeRIC-5>> at 24 April 2008.

³⁰ <<http://imos.org.au/about.html>> at 20 May 2008.

³¹ <<http://imos.org.au>> at 20 May 2008.

Sea and the Great Barrier Reef.³² Among other things, GBROOS will monitor the effect of rising ocean temperatures on the incidence of coral bleaching over the next decade.³³ GBROOS includes the world's first large scale reef-based Internet Protocol (IP) network.³⁴ Data generated by the IMOS project will be made available to researchers through the electronic Marine Information Infrastructure (eMII) located at the University of Tasmania.³⁵ After defining specific data streams, IMOS will eventually develop end-to-end protocols, standards and systems to join the related observing systems into a unified data storage and access framework.³⁶ Data will be archived within the Australian Ocean Data Network (AODN), which is a distributed data storage and discovery network based at leading Australian marine research facilities.³⁷ Data storage and retrieval in IMOS is designed to be interoperable with other national and international programs.³⁸ IMOS is an NCRIS funded project.³⁹

Another example is the Pacific and Regional Archive for Digital Sources in Endangered Cultures (PARADISEC).⁴⁰ PARADISEC offers a facility for digital conservation and access for endangered materials from the Pacific region, defined broadly to include Oceania and East and Southeast Asia.⁴¹ PARADISEC is also a national repository for recorded material relating to indigenous cultures of regions in and around Australia. PARADISEC has established a framework for accessing, cataloguing and digitising audio, text and visual material, and preserving digital copies.⁴² The project has been funded by the

³² <[http://imos.org.au/newsitem.html?&tx_ttnews\[tt_news\]=64&tx_ttnews\[backPid\]=2&cHash=d32f9070cb](http://imos.org.au/newsitem.html?&tx_ttnews[tt_news]=64&tx_ttnews[backPid]=2&cHash=d32f9070cb)> at 20 May 2008.

³³ <<http://imos.org.au/gbroos.html>> at 20 May 2008.

³⁴ <[http://imos.org.au/newsitem.html?&tx_ttnews\[tt_news\]=64&tx_ttnews\[backPid\]=2&cHash=d32f9070cb](http://imos.org.au/newsitem.html?&tx_ttnews[tt_news]=64&tx_ttnews[backPid]=2&cHash=d32f9070cb)> at 20 May 2008.

³⁵ <<http://imos.org.au/about.html>> at 20 May 2008, see also <<http://imos.org.au/emii.html>> at 20 May 2008.

³⁶ <<http://imos.org.au/emii.html>> at 20 May 2008.

³⁷ <http://imos.org.au/data_access.html> at 20 May 2008.

³⁸ <<http://imos.org.au/emii.html>> at 20 May 2008.

³⁹ <<http://imos.org.au/about.html>> at 20 May 2008.

⁴⁰ <<http://paradisec.org.au/home.html>> at 25 April 2008.

⁴¹ <<http://www.paradisec.org.au/about.html>> at 25 April 2008.

⁴² <<http://www.paradisec.org.au/about.html>> at 25 April 2008.

Universities of Sydney, Melbourne and New England, ANU, the Australian Research Council (ARC) and GrangeNet.⁴³

Funding policies

Australian funding bodies have taken an interest in the management of and provision of access to research data. In December 2006, the Australian Research Council (ARC) and the National Health and Medical Research Council (NHMRC) announced the introduction of open access guidelines for published papers and data resulting from funded research projects, effective 2008. Both policies encouraged researchers to:

Consider the benefits of depositing their data and any publications arising from a research project in an appropriate subject and/or institutional repository [because in order to] maximise the benefits from research, findings need to be disseminated as broadly as possible to allow access by other researchers and the wider community.⁴⁴

The same guidelines are contained in the ARC *Discovery Project Funding Rules for funding commencing in 2009*,⁴⁵ and the NHMRC *Project Grants Funding Policy for funding commencing in 2009*.⁴⁶

The introduction of open access requirements for data resulting from funded research projects was supported by the Australian Government Productivity Commission in its 2007 report, *Public Support for Science and Innovation*.⁴⁷ The Productivity Commission commended the steps taken

⁴³ <<http://www.paradisec.org.au/about.html>> at 25 April 2008.

⁴⁴ Australian Research Council (ARC), *Discovery Projects Funding Rules for funding commencing in 2008*, [1.4.5.1] and [1.4.5.3] <http://www.arc.gov.au/pdf/DP08_FundingRules.pdf> at 25 April 2008; National Health and Medical Research Council (NHMRC), *Project Grants Funding Policy for grants commencing in 2008*, [16.2]. See also A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 4.

⁴⁵ Australian Research Council (ARC), *Discovery Projects Funding Rules for funding commencing in 2009*, [4.4.5.1] and [4.4.5.3] <http://www.arc.gov.au/ncgp/dp/dp_fundingrules.htm> at 25 March 2008.

⁴⁶ National Health and Medical Research Council (NHMRC), *Project Grants Funding Policy for funding commencing in 2009*, [16.2] <<http://www.nhmrc.gov.au/FUNDING/apply/granttype/projects/index.htm>> at 25 March 2008.

⁴⁷ Productivity Commission, *Public Support for Research and Innovation*, Research Report (2007) 240, 243 <<http://www.pc.gov.au/study/science/docs/finalreport>> at 25 April 2008.

by the ARC and NHMRC to promote open access to the results of the projects they fund. However, the Productivity Commission considered that in light of experience in the United States voluntary compliance was likely to be low. Consequently, the Productivity Commission considered that the aim of free and open access to publicly-funded research results would be better achieved by the progressive introduction of mandatory open access requirements.⁴⁸

Surveys of researchers

While the benefits of data sharing have been widely recognised by government agencies and scientific organisations, there is a degree of reluctance among researchers to embrace data sharing practices. Recent surveys of the Australian research community provide indications of current attitudes and practices in relation to data ownership and sharing.

The NCRIS Platforms for Collaboration Data Management Survey conducted in September and October 2006 surveyed key stakeholders in the management of research data throughout Australia.⁴⁹ The results of the survey demonstrated that while some researchers are aware of the complexity of the issues involved in data ownership, most have only a rudimentary understanding.⁵⁰ Further, the survey made clear that while there is an awareness of the potential benefits of data sharing within the Australian research community, there are also concerns about the exploitation of data by others, especially if this would diminish the credit attributed to the researcher who generated the data in the first place.⁵¹ The NCRIS survey made apparent the need for researchers to be provided with guidelines and data management infrastructure to assist in developing a better understanding of data ownership and management.⁵²

In October 2006, the Australian Partnership for Sustainable Repositories (APSR) project published the results of a survey of researchers based

⁴⁸ Productivity Commission, *Public Support for Research and Innovation*, Research Report (2007) 240, 243 <<http://www.pc.gov.au/study/science/docs/finalreport>> at 25 April 2008. See also A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 4.

⁴⁹ See <www.pfc.org.au/twiki/pub/Main/DataWorkshop1/NCRISsurveyanalysis.pdf> at 25 April 2008.

⁵⁰ See A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 128.

⁵¹ See A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 128.

⁵² See A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 128.

across several research institutions and research service providers throughout Australia. The results of this Australian e-Research Sustainability Survey (AERES) were published in a report entitled, *Sustainability Issues for Australian Research Data: the report of the Australian e-Research Sustainability Survey Project*.⁵³ The AERES study found a distinct lack of formal policies for data management utilised by the surveyed researchers.⁵⁴ The report concluded that current data practices generally see data managed sufficiently for research needs but not professionally; discoverable through scientific publication but not otherwise; and having a value placed on it for present needs but not for the future.⁵⁵

In August 2007, the Legal Framework for e-Research Project based at the Queensland University of Technology (QUT) published the report, *Legal and project agreement issues in collaboration and e-Research: Survey Results*.⁵⁶ This report documented a survey that was conducted online during May 2007 and was open to all Australian participants involved in collaborative research.⁵⁷ The QUT survey found that many researchers consider legal agreements to be an impediment to timely research and will often commence collaborative research projects before finalising agreements dealing with data ownership and other legal interests.⁵⁸ One participant, a university researcher in the Arts and Social Sciences, responded, 'Perhaps the biggest problem facing e-Research is the lack of understanding and agreement as to what is required in terms of local and

⁵³ Markus Buchhorn and Paul McNamara, *Sustainability Issues for Australian Research Data: The Report of the Australian e-Research Sustainability Survey Project*, Australian Partnership for Sustainable Repositories (APSR) (2006) <<http://www.apsr.edu.au/aeres>> at 25 April 2008.

⁵⁴ A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 126.

⁵⁵ Markus Buchhorn and Paul McNamara, *Sustainability Issues for Australian Research Data: The Report of the Australian e-Research Sustainability Survey Project*, Australian Partnership for Sustainable Repositories (APSR) (2006) 44 <<http://www.apsr.edu.au/aeres>> at 25 April 2008.

⁵⁶ Maree Heffernan and Nikki David, *Legal and project agreement issues in collaboration and e-Research: Survey Results*, Legal Framework for e-Research Project, Queensland University of Technology (QUT) (2007) <<http://www.e-research.law.qut.edu.au/>> at 25 April 2008.

⁵⁷ <<http://www.e-research.law.qut.edu.au/>> at 25 April 2008.

⁵⁸ See Maree Heffernan and Nikki David, *Legal and project agreement issues in collaboration and e-Research: Survey Results*, Legal Framework for e-Research Project, Queensland University of Technology (QUT) (2007) 38, 62 <<http://www.e-research.law.qut.edu.au/>> at 25 April 2008.

national information infrastructure to support e-Research activities'.⁵⁹ The QUT survey highlighted the need for simple and easy-to-use resources to assist researchers in managing the legal rights surrounding data and e-Research, particularly where collaborative research projects are concerned.

LEGAL IMPLICATIONS SURROUNDING DATA ACCESS, SHARING AND REUSE

The collection, management and use of research data occurs in a legal context and raises a host of legal issues. Quite simply, data is surrounded by law.⁶⁰ For example, arrangements between a researcher and other researchers, research institutions or funding bodies may be governed by contract. Data compilations may attract copyright protection and data may also attract protection under confidentiality or privacy laws. This section provides an overview of the different legal regimes that may apply to and impact upon data collection, access, sharing and reuse.

Copyright

A general principle of copyright law is that copyright protects the material form in which ideas, information or facts are expressed and not the ideas, information or facts themselves. It follows that under this general principle, copyright law will not protect raw data. However, in Australia, copyright law may operate to protect compilations of data, such as datasets or databases, provided that the compilation meets the originality threshold required by law. Under the *Copyright Act 1968* (Cth), a compilation is protected as a literary work.⁶¹

Compiled data will not always be raw data – a compilation may also include written materials, reports, diagrams, tables and graphs. Where a data item meets the form and originality requirements under the *Copyright Act*, it may be protected by copyright as an independent work.

⁵⁹ Marce Heffernan and Nikki David, *Legal and project agreement issues in collaboration and e-Research: Survey Results*, Legal Framework for e-Research Project, Queensland University of Technology (QUT) (2007) 62 <<http://www.e-research.law.qut.edu.au/>> at 25 April 2008.

⁶⁰ A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 263.

⁶¹ *Copyright Act 1968* (Cth) s 10(1).

An important distinction lies between copyright in discrete data items and copyright in a database as a whole. In the latter, copyright serves to protect the *arrangement* of the collected components. Copyright interests may co-exist independently in components contained within the database and in the database itself, and may be owned by different parties.⁶²

Copyright will only protect a work that possesses the requisite level of originality under law. In *Desktop Marketing v Telstra*,⁶³ the court considered the issue of whether a compilation is sufficiently original to attract copyright protection. The question for the court was whether Telstra held copyright in their White Pages and Yellow Pages directories, which are essentially a compilation of names, addresses and phone numbers listed alphabetically. In a landmark judgment, the court held that Telstra did own copyright in their compilations, thereby establishing that the originality threshold for copyright protection is low. The court held that copyright can be claimed in a compilation that:

1. has been produced as a result of the exercise of skill, judgment or knowledge in the selection, presentation or arrangement of the materials; or
2. has required the investment of a substantial amount of labour or expense to generate or collect the material included in the compilation (the so-called ‘sweat of the brow’ approach).⁶⁴

Telstra, in undertaking substantial labour and incurring substantial expense, had met the originality threshold in compiling the Yellow Pages and White Pages directories, notwithstanding that there may have been minimal intellectual input or creativity involved in the selection and arrangement of the material.

Significantly, the court in this decision prescribed a lower threshold for originality under Australian copyright law than that required in the

⁶² A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 137.

⁶³ *Desktop Marketing Systems Pty Ltd v Telstra Corporation Ltd* [2002] FCAFC 112. See also *Nine Network Australia Pty Ltd v Ice TV Pty Ltd* [2008] FCAFC 71.

⁶⁴ *Desktop Marketing Systems Pty Ltd v Telstra Corporation Ltd* [2002] FCAFC 112, [409].

United States, where there must be a degree of creativity applied in the selection, coordination or arrangement of the compilation.⁶⁵

The owner of copyright in a database, dataset or where applicable, a discreet item of data will be able to control how that database, dataset or data is used, copied and shared. It would be wise for a research project that intends to allow its data to be openly shared and reused to formulate plans and policies that properly define, allocate and manage copyright interests in the data and database.

Patents

Patents protect products and processes that are novel, useful and involve an inventive or innovative step.⁶⁶ They confer on the patentee the exclusive right to exploit the patented product or process for a period of time (usually 20 years from the time of filing the patent application).⁶⁷ Data or information can be practically applied in such a way that it forms part of or gives rise to an invention capable of being patented. This situation has most commonly arisen in the context of patenting genomic data.⁶⁸

Researchers collecting data may be concerned with patents for one of two reasons. Firstly, some researchers may be interested in obtaining a patent over a product or process that incorporates data which they have collected. For these researchers, disclosure of data could prevent a patent being obtained because releasing information into the public domain could preclude the ‘novel’ or ‘inventive’ aspect of a product or process that is required by law to secure a patent.⁶⁹ In these circumstances, prior to obtaining a patent, data should only be disclosed under confidentiality agreements to ensure that the data is kept out of the public domain.

Secondly, some researchers may want to ensure that their data is kept free of legal restrictions including patents, in order to allow sharing and

⁶⁵ See *Feist Publications Inc v Rural Telephone Service Co Inc*, 499 US 340, 349 (1991).

⁶⁶ *Patents Act 1990* (Cth) s 18.

⁶⁷ *Patents Act 1990* (Cth) ss 13 and 67.

⁶⁸ For more information, see A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 116–23.

⁶⁹ See *Patents Act 1990* (Cth) s 7.

reuse by themselves and others. For these researchers, simply releasing data into the public domain may be enough to create prior art and thus prevent successful patent applications by others.⁷⁰ However, even where data is released publicly it may be possible for another party to make improvements to the disclosed data and then make these improvements proprietary. Where data is used to develop a patentable invention, the subsequent patent rights may be broad enough to cover use of the actual data forming part of the invention.⁷¹ Fortunately, there are contractual and licensing options that can be employed to keep data free of restrictive patent claims. One option is to release data via an online database where users accessing the database are required to first enter into a click-wrap agreement that governs use of the data. The agreement can prohibit patent applications based on certain data, or may allow patent applications but provide that the patent must not be restrictive and must allow further use of the patented data.⁷² Another option is to actually obtain a patent over a product or process based on or encompassing the research data, but then to licence the use of the protected data under liberal terms.⁷³

⁷⁰ This was the approach underlying the Bermuda Principles, which were developed in 1996 by a consortium of researchers involved in the Human Genome Project. For more information, see A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 118–20.

⁷¹ See A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 119–20; Donna M Gitter, 'Resolving the Open Source Paradox in Biotechnology: A Proposal for a Revised Open Source Policy for Publicly Funded Genomic Databases' (2007) 43(4) *Houston Law Review* 4 <<http://ssrn.com/abstract=901994>>; Rebecca Eisenberg and Arti Rai, 'Harnessing and Sharing the Benefits of State-Sponsored Research: IP Rights and Data Sharing in California's Stem Cell Initiative' (2006) 21 *Berkeley Technology Law Journal* 1187, 1207; Claire T Driscoll, 'NIH data and resource sharing, data release and intellectual property policies for genomics community resource projects' (2005) 15(1) *Expert Opinion on Therapeutic Patents* 4.

⁷² This was the approach adopted by the International Haplotype Project (commonly known as the HapMap Project), which ran from 2002 to 2005. For more information, see A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 120–21.

⁷³ This was the approach adopted by the CAMBIA project. See CAMBIA, 'About BiOS (Biological Open Source) Licenses and MTAs' <<http://www.bios.net/daisy/bios/licenses/398.html>> at 11 April 2008. See also, Richard Jefferson, 'Science as Social Enterprise: The CAMBIA BiOS Initiative' (2006) 1(4) *Innovations: Technology, Governance, Globalization* 13; and A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 121–3.

Privacy

Some research, particularly research in medical fields, will give rise to privacy concerns about the handling and use of personally identifying and health information. In response to privacy concerns, the *Privacy Act 1988* (Cth) requires Commonwealth public sector entities to act in accordance with Information Privacy Principles and private sector entities to act in accordance with National Privacy Principles. The Information Privacy Principles prevent the collection of personal information by a government agency except where the collection is for a lawful purpose directly related to a function or activity of the agency.⁷⁴ The National Privacy Principles provide that personal information cannot be used except for the lawful purpose for which it was collected.⁷⁵ ‘Personal information’ is defined in the *Privacy Act* as ‘information or an opinion ... about an individual whose identity is apparent, or can reasonably be ascertained, from the information or opinion.’⁷⁶

The National Privacy Principles differentiate between ‘personal information’ and ‘sensitive information’. Sensitive information is accorded a higher level of protection and is defined to include health or genetic information about an individual.⁷⁷ An organisation must generally not collect sensitive information about an individual unless the individual has consented.⁷⁸ Obtaining consent to collect sensitive information for research purposes will usually involve explaining to the participant the purpose, methods, possible risks and potential outcomes of the research, including the likelihood that research results will be published.

There are limited exceptions to the requirements imposed in the Information Privacy Principles and the National Privacy Principles relating to the collection and disclosure of personal and sensitive information. For example, consent to disclose personal information will not be required where the participant was reasonably likely to have been

⁷⁴ *Privacy Act 1988* (Cth) s 14: Information Privacy Principle 1.

⁷⁵ *Privacy Act 1988* (Cth) Schedule 3.

⁷⁶ *Privacy Act 1988* (Cth) s 6.

⁷⁷ *Privacy Act 1988* (Cth) s 6; Schedule 3: National Privacy Principle 10.

⁷⁸ *Privacy Act 1988* (Cth) Schedule 3: National Privacy Principle 10.

aware or would reasonably expect that the information would be disclosed, or where it is impractical to obtain consent.⁷⁹

The definition of ‘personal information’ refers to information that can be used to identify an individual. Where information has been de-identified such that it cannot be re-identified, it can usually be used and disclosed in research and data-linkage without fear of infringing the *Privacy Act*. Studies show that individuals generally support the idea of researchers being able to access health information from databases, provided that the information is identified by a unique number rather than a name.⁸⁰ The National Health and Medical Research Council (NHMRC) has recommended the transitory use of patient identifiers for the purposes of data-linkage, even without patient consent, provided that the personal information enabling linkage is not retained after the linkage, the identifying information is used with sufficient security and the research for which the data is being linked has public benefit.⁸¹

All Australian States and Territories except Queensland and South Australia have enacted privacy legislation or introduced privacy bills relating to health information and/or the collection and use of personal information in the State public sector.⁸² In South Australia, the Privacy Committee is responsible for administrative protocol PC012 – Information Privacy Principles Instruction, which applies to public

⁷⁹ See *Privacy Act 1988* (Cth) s 14; Information Privacy Principle 11.1; Schedule 3: National Privacy Principles 2.1 and 10.3; Australian Law Reform Commission, *Review of Privacy* (Issues Paper 31, 2006) [8.124].

⁸⁰ Australian Law Reform Commission, *Review of Privacy* (Issues Paper 31, 2006) [8.237].

⁸¹ National Health and Medical Research Council (NHMRC), *National Statement on Ethical Conduct in Research Involving Humans* (1999) 53 <http://www.nhmrc.gov.au/publications/synopses/_files/e35.pdf>. Note: the 2007 revised National Statement on Ethical Conduct in Research Involving Humans was tabled in Parliament on 28 March 2007.

⁸² *Privacy and Personal Information Act 1998* (NSW), *Health Records and Information Privacy Act 2002* (NSW), *Information Privacy Act 2000* (Vic), *Health Records Act 2001* (Vic), *Australian Capital Territory Government Service (Consequential Provisions) Act 1994* (Cth), *Health Records (Privacy and Access) Act 1997* (ACT), *Information Act 2002* (NT), *Personal Information Protection Act 2004* (Tas), *Health Complaints Act 1995* (Tas). For Western Australia, see *State Records Act 2000* (WA) s 49 and *Freedom of Information Act 1992* (WA) s 3. In March 2007, the *Information Privacy Bill 2007* (WA) was introduced into the Legislative Assembly. It passed the Legislative Assembly on 27 November 2007 and reached the second reading speech stage in the Legislative Council on 4 December 2007 (see <<http://www.parliament.wa.gov.au/web/newwebparl.nsf/iframewebpages/Bills+-+Current>> at 20 May 2008).

sector handling of personal information. In Queensland, there are two administrative protocols applying to the State's public sector. Information Standard 42: Information Privacy applies to the collection of personal information in the public sector generally, while Information Standard 42A: Information Privacy for the Queensland Department of Health applies to the collection of health information. The differences in privacy regulation at Commonwealth and State levels has caused some confusion for medical researchers, prompting the NHMRC, the Australian Government Productivity Commission and the Australian Law Reform Commission (ALRC) to recommend a nationally consistent approach to privacy regulation of health information.⁸³

Confidential Information

Data that has not been released into the public domain may be protected by the law of confidentiality. A researcher who has expended considerable time and energy in generating or collecting data may have an interest in protecting that the data from others who have not contributed to its production. In such a situation, the action for breach of confidence can be used to control access to the data.

The law of confidentiality is based on the equitable principle that a person who receives information in confidence shall not take unfair advantage of that information.⁸⁴ A successful breach of confidence action must establish three elements:

1. the information is confidential in nature;
2. the information was imparted in circumstances importing an obligation of confidence; and
3. an unauthorised use of the information to the detriment of the person claiming the right to maintain confidentiality.⁸⁵

Data will only be protected as confidential if it is not in the public domain. A breach of confidence action can still be established where

⁸³ Australian Government Productivity Commission, *Public Support for Science and Innovation*, Research Report (2007) 189, 217; Australian Law Reform Commission, *Review of Australian Privacy Law: An Overview of Discussion Paper 72* (2007) 5–6.

⁸⁴ *Seager v Copydex Ltd* [1967] 2 All ER 415, 417.

⁸⁵ *Coco v A N Clark (Engineers) Ltd* [1969] RPC 41, 47.

more than one person knows about or has access to the data, provided that not so many people know about the data that it can no longer be regarded as secret. Usually, confidentiality will be protected through the use of confidentiality agreements, which provide for the disclosure of information on the condition that the contracting party does not further disclose the information and does not use the information except for the purposes set out in the agreement.

Confidentiality will be lost if enough people know about the data such that it passes into the public domain, or if the data is independently discovered by someone else.⁸⁶ Where data loses its quality of secrecy, it is still possible for a researcher to control access to and use of the data through contract.

Contract

In practice, the most important legal mechanism used to allocate rights to data is the contractual agreement. There are three main types of agreement relevant to regulating data access and use – the confidentiality agreement, the copyright licence and the access agreement.

Confidentiality agreements, also called non-disclosure agreements, serve to protect secret information by disclosing the information in a controlled setting so that it remains confidential and is not released into the public domain. Confidentiality agreements will generally: identify the owner of rights in relation to the confidential information; identify the information that is to be treated as confidential; impose obligations on the person to whom the information is disclosed to maintain the secrecy of the information; define the scope of the permitted use of the information; and provide for the consequences of a failure to comply with the confidentiality obligations.⁸⁷

Copyright licences grant permission to a person to deal with a database or a dataset in a way that would otherwise infringe copyright. For example, a copyright owner may permit - through a licence - a researcher to reproduce copyright material and make the material available on a website where it can be accessed and downloaded by other researchers. A contractual copyright licence may also contain terms that

⁸⁶ *Attorney General (UK) v Heinemann Publishers Australia Pty Ltd* (1988) 10 IPR 153.

⁸⁷ A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 175.

are not strictly related to copyright. For example, the licence may require the researcher to undertake not to hold the copyright owner liable for consequences resulting from any inaccuracies that may be contained in the data supplied.⁸⁸ Contractual licences will usually indicate the copyright material to which the licence refers; the permitted acts that the licensee is authorised to do; any restrictions upon the party acting under the licence; the consideration provided for the licence; and whether or not the licence is exclusive (or non-exclusive) and whether it can be revoked or is irrevocable.⁸⁹

Access agreements will operate where a researcher or research organisation has control over the database in which their data is stored. The researcher or research organisation may require persons interested in obtaining access to the data to first enter into an access agreement. Access agreements may: identify the data to be accessed; identify the person/s or class of persons who are permitted to access the data; state that access rights cannot be transferred to third parties; limit the purposes for which the data may be used; contain a disclaimer that the researcher is not responsible for any inaccuracies in the data; and provide for the consequences of a failure to comply with the agreement.⁹⁰ For example, an access agreement may provide that the data can be accessed and used for non-commercial purposes only, or may provide that if a user engages in commercial uses of the data, they must account back to the researcher for a proportion of the profits. Access agreements can be used to control access to and use of data that was formerly protected through confidentiality agreements but which has lost its quality of confidence.

DATA SHARING INFRASTRUCTURE

It will not be sufficient for researchers and database managers to simply be aware of the laws that surround the data they collect. If data is to be effectively made available within the research community, it is necessary that it is properly managed. Research projects would be wise to adopt protocols for dealing with the legal issues that may arise in relation to the

⁸⁸ A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 173.

⁸⁹ A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 177.

⁹⁰ A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 176–7.

data they collect. Failure to establish legal protocols for data management may jeopardise the research community's ability to access, share and use valuable research outputs.⁹¹ Data sharing infrastructure, such as data management policies, principles, plans and toolkits, can assist researchers and database managers to effectively manage their legal rights, interests and obligations in relation to the data collected, generated and compiled by the research project.

Data management policies and principles

A forward-thinking research project will have in place a data management policy containing high-level statements about how data generated or compiled by the research project is to be made available for access and use. The data management policy may also contain principles expanding on the high-level statements and indicating how they are to be applied.⁹²

A data management policy will take into consideration the research discipline of the project; the funding arrangements for the research project; the kind of data generated or collected by the project; how and when data is to be deposited into a database; how, when and on what basis data is to be made available for access by other researchers; and any legal obligations imposed on the research project or individual researchers.⁹³

A research project must give careful consideration to formulating a policy which ensures that researchers' objectives, needs and responsibilities in each research situation are properly addressed.⁹⁴ For example, where a research project is publicly funded, it may be appropriate for a policy to strongly support immediate open access to research data. However, immediate open access may not be appropriate for data generated by private sector research projects.

The Australian Partnership for Sustainable Repositories (APSR) has highlighted the importance of all data management policies including

⁹¹ A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 263.

⁹² A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 240.

⁹³ A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 241.

⁹⁴ A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 241.

clear definitions of concepts and terms used within the policy.⁹⁵ Additionally, research projects should take care to distinguish in their policies data that is to be made accessible from data that is not. This is particularly important where a research project is collecting data subject to privacy limitations or data that is to be commercially exploited.

Data management policies and principles will also explain the conditions under which data is to be made available for access and use. For example, access may be limited to certain categories of researchers or researchers may only be permitted to use the data for specified purposes. In order to properly ascertain and set out the conditions of access and use, each research project should develop a clear and comprehensive listing of all legal restrictions applying to the management, dissemination and reuse of the different kinds of data that may be generated by the project.⁹⁶

Data management plans

Similar to a data management policy, a data management plan (DMP) will address how data is collected, stored, managed and disseminated. It will also be concerned with data ownership and the legal controls surrounding data. However, a DMP will focus on practical measures rather than making broad policy statements. It will also consider expenditures and technical measures to ensure sustainability of data.⁹⁷

A DMP should be in place from the conception and commencement of a research project. A comprehensive DMP will recognise that there are many different parties involved in a research project and will have relevance to all of the different parties. These parties include collectors and compilers of data, data analysts, database managers, parties that have funded the research project and consumers or users of the data and database.⁹⁸

⁹⁵ Anna Shadbolt et al, *Sustainable Paths for Data-intensive Research Communities at the University of Melbourne: A Report for the Australian Partnership for Sustainable Repositories* (August 2006) 38–9 <http://www.apsr.edu.au/acres/sustainable_paths.pdf>.

⁹⁶ A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 244.

⁹⁷ A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 247–56.

⁹⁸ A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 247–8.

It is important that a DMP addresses unusual situations that may arise in the collation of data. For example, where data that is generated by the research project is to be integrated with existing data from other sources, the DMP will need to explain how this will be done and how data from each source will be identified once combined. It must also ensure that legal rights and obligations are respected.

Two central issues for each research project to consider in its DMP are:

1. who owns the data generated or collected by the research project; and
2. who is responsible for managing the data?

Data may be owned by more than one person. An owner may be the researcher who has collected or generated the data; the researcher's employer, under the terms of the researcher's employment contract; the funder of the research, under the terms of the funding agreement; or the database owner or provider. Each party's ownership rights will need to be defined in the DMP. Additionally, the DMP should set out who is responsible for managing the data. Management responsibilities may include recording, organising and archiving the data and managing access to the data. A comprehensive DMP will address the management roles of each party and will set out the formal levels of responsibility required for database management and maintenance.⁹⁹

As explained above, data collection, access and reuse will be affected by legal controls. It is imperative that a DMP considers the legal and regulatory controls applying to the data that is collected by the research project. Such legal controls may include confidentiality restrictions for secret information, copyright assignments and licences, deposit agreements for inclusion of data in a database and agreements governing access to that database. All contractual obligations should be considered and addressed. In particular, a DMP should describe the conditions under which the research project is funded and any obligations – contractual or otherwise – that the researchers have to the funding body. Finally, a DMP should consider whether legislation applies to the collection or use of data, such as the application of privacy legislation for projects dealing with personal information.¹⁰⁰

⁹⁹ A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 250, 254.

¹⁰⁰ A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 251–2.

Data security and sustainability are two important considerations for any DMP. The level of security that will operate in relation to the data collected will vary depending on the type of data concerned. For example, more stringent security may be applied to data that is confidential or which may form the basis of a patent application. For these types of data, access may be limited to select individuals (access may be password protected) and reuse rights may be minimal. Contractual agreements may regulate what disclosures can and cannot be made in relation to the data. For less sensitive data, the applicable security measures are likely to be less strict. A DMP will need to set out the different security measures relevant to the different levels of data and how these security measures are to be implemented.¹⁰¹

Careful consideration must be given to the potential future relevance of any data collected or generated by the research project. Where it is envisaged that data could be useful for future research, sustainability of data will be an important issue to address. A DMP should describe whether long-term preservation of the data is necessary and if so, how long the data will be preserved and who will be responsible for ensuring its preservation. A related issue will be how to ensure the ongoing, long-term funding of the database even after the research project that gave rise to the database is finished.¹⁰²

Data management toolkits

A data management toolkit (DMT) is a document aimed at researchers within a research project, which provides practical guidelines about implementing the DMP. A DMT can assist individual researchers in ascertaining their role and level of responsibility within a research project and with understanding what is to be done with the data collected or generated by the project. A DMT can inform researchers about who will be able access the data collected by the researchers and how they may reuse that data. It can also assist researchers in determining their obligations, both legal and otherwise, in relation to the data that they generate or collect.

¹⁰¹ A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 252–4.

¹⁰² A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 255.

A DMT can be tailored to different levels of research and researchers. It may be appropriate to have a different DMT applying to researchers than that applying to database managers, or a different DMT applying to a small research team within a single institution than that applying to a larger research team that is part of a collaborative project spread across many institutions.

A DMT provides practical guidance to assist researchers in managing their data in compliance with the project's data management policies and procedures, DMP and the relevant legal framework. Therefore, a DMT should take the form most accessible to a project's researchers, whether this be in the form of a textual document, a series of questions, diagrams or multimedia tools. Yet regardless of form, all DMTs should enable researchers to understand the ownership and management issues surrounding data collection and compilation; the legal and technical restraints applying to collection, storage, handling and use of data; and the access, sharing, use and reuse framework surrounding the project's data.¹⁰³

Licensing models

As far as data and databases attract copyright, licences can be used to allow access to and reuse of the data by other researchers. The emergence of open content licensing models has made it much easier for copyright owners to licence their material to a wider range of people, especially where it is distributed over the internet.¹⁰⁴ Open content licensing involves making copyright material available on liberal terms, to ensure that it is readily accessible and available for reuse.¹⁰⁵ The last few years have seen an increasing appreciation of open content licences to grant access to copyright-protected data collections in open collaborative research projects.¹⁰⁶

¹⁰³ See A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 256–7.

¹⁰⁴ A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 146.

¹⁰⁵ A Fitzgerald and K Pappalardo, *Building the Infrastructure* (2007) 146.

¹⁰⁶ A Fitzgerald and K Pappalardo, *Building the Infrastructure*, 148. See for example, Editorial, 'Let data speak to data' (2005) 438 *Nature* 531 <<http://www.nature.com/nature/journal/v438/n7068/full/438531a.html>>; Don Tapscott and Anthony Williams, 'The New Science of Sharing' (March 2007) *BusinessWeek.com* <http://www.businessweek.com/innovate/content/mar2007/id20070302_219704.htm?chan=technology_technology+index+page_more+of+today's+top+stories>;

The leading model of open content licensing is the suite of Creative Commons licences developed by the Creative Commons Project.¹⁰⁷ The Creative Commons (CC) Licences make copyright works freely available for use, on certain conditions as selected by the licensor. Where one or more elements of a database attracts copyright, a CC licence can be used to licence that copyright to users. For example, the CC licensing model is utilised by the Universal Protein Resource (UniProt), a comprehensive resource for protein sequence and annotation data and a collaboration between the European Bioinformatics Institute (EBI), the Swiss Institute of Bioinformatics (SIB) and the Protein Information Resource (EBI).¹⁰⁸ UniProt has chosen to apply the Creative Commons Attribution-No Derivatives Licence to all copyrightable parts of its databases.¹⁰⁹

Science Commons is a project related to Creative Commons that extends open access principles to scientific data and publications.¹¹⁰ Formerly, the Science Commons 'Databases and Creative Commons FAQ' stated that a CC licence could be applied to copyrightable elements of a database, but advised database providers to:

- understand and make clear on the database website which elements of the database are licensed under the CC licence, based on the existence of copyright in those elements;
- understand and make clear on the database website which parts of the database are not subject to copyright (ie raw data and information) and which are therefore free to be used and reused independently of the CC licence;
- ensure that they have the necessary authority to apply a CC licence to the database (ie that they are the copyright owner or have permission from the copyright owner);
- where applicable, inform users that the CC licence only applies to the database elements and not the underlying software; and

Charlotte Waelde and Mags McGinley, 'Designing a licensing strategy for sharing and re-use of geospatial data in the academic sector' (2007) GRADE <<http://edina.ac.uk/projects/grade>>.

¹⁰⁷ See <<http://www.creativecommons.org>> or <<http://www.creativecommons.org.au>>.

¹⁰⁸ 'About UniProt' <<http://beta.uniprot.org/help/about>> at 22 April 2008.

¹⁰⁹ *License & disclaimer* <<http://beta.uniprot.org/help/license>> at 22 April 2008.

¹¹⁰ See <<http://sciencecommons.org>>.

- be aware that CC licences do not licence all types of legal rights, but only licence copyright, and so legal restrictions relating to patents, privacy, confidentiality, contract and other relevant legal frameworks will not be affected by the adoption of a CC licence.¹¹¹

Science Commons has since moved away from endorsing the application of CC licences to databases. The recommendation has been withdrawn because of difficulties identified by Science Commons with:

1. identifying the copyrightable and non-copyrightable elements of a database, such that obligations based on copyright (eg the option under some CC licences to require that use of the copyright material is non-commercial) are imposed in situations where copyright does not apply and the obligation is inappropriate; and
2. proper attribution (a requirement under all CC licences), where hundreds or even thousands of scientists have potentially contributed to or deposited data in the database.¹¹²

Science Commons has instead developed a 'Protocol for Implementing Open Access Data', which sets out the principles for open access data and provides a protocol for implementing those principles. Additionally, Science Commons distributes an Open Access Data Mark and metadata for use on data and databases that conform to the protocol. The protocol is not a copyright licensing model. Instead, the protocol requires a waiver of legal rights and all legal grounds for database protection in order to dedicate the data to the public domain.¹¹³ Science Commons acknowledges that the protocol will not be appropriate for all

¹¹¹ See <<http://sciencecommons.org/resources/faq/databases/>> at 22 April 2008.

¹¹² Science Commons, *Protocol for Implementing Open Access Data* <<http://sciencecommons.org/projects/publishing/open-access-data-protocol/>> at 22 April 2008. See also Science Commons, *Database Protocol FAQ* <<http://sciencecommons.org/projects/publishing/open-access-data-protocol/>> at 22 April 2008.

¹¹³ Science Commons, *Protocol for Implementing Open Access Data* <<http://sciencecommons.org/projects/publishing/open-access-data-protocol/>> at 22 April 2008. See also Science Commons, *Database Protocol FAQ* <<http://sciencecommons.org/projects/publishing/open-access-data-protocol/>> at 22 April 2008.

types of data, but believes that the protocol offers a system that is both legally accurate and easier for scientists to understand than many copyright licensing models.¹¹⁴

The problems identified by Science Commons are indeed apparent in many situations involving licensing of database elements and are worthy of careful consideration. However, the legal position regarding copyright protection of data and databases is much more straightforward in Australia than in either the United States or Europe. It is considerably easier to distinguish between the copyright and non-copyright elements of databases in Australia than in the United States, where the creativity of a compilation must be assessed before copyright applies. Further, there are fewer legal considerations in Australian than in Europe, where a *sui generis* database right operates to protect databases irrespective of whether the database or its contents attract copyright protection.

The concerns raised by Science Commons highlight the importance of each and every research project adopting a DMP that properly considers and manages issues of data ownership and legal rights including copyright. It is entirely possible to successfully apply a copyright licensing model to a database and its copyrightable contents. However, in order to ensure the successful operation of the licence, the research project's DMP must clearly identify which legal rights apply to which database elements and which database elements are to be licensed to the public and on what terms. The DMP should also state how the data and database are to be attributed and make this information readily apparent on the database website. For example, UniProt provides a webpage that informs users how to cite resources and publications obtained from the UniProt website or databases under a CC licence.¹¹⁵

¹¹⁴ Science Commons, *Protocol for Implementing Open Access Data* <<http://sciencecommons.org/projects/publishing/open-access-data-protocol/>> at 22 April 2008. See also Science Commons, *Database Protocol FAQ* <<http://sciencecommons.org/projects/publishing/open-access-data-protocol/>> at 22 April 2008.

¹¹⁵ *Publications – How to cite us* <<http://beta.uniprot.org/help/publications>> at 22 April 2008.

CONCLUSION

For any research project, several important legal and management decisions will need to be made about the data collected or generated in the course of the research. How will ownership interests in data be determined and allocated? Will data be made accessible to the public, and if so, on what basis? Will sharing and reuse of data be permitted? What legal restraints apply to the data? All these questions must be carefully considered, answered and agreed upon by members of the research project, including researchers, database managers, hosting institutions and funding bodies.

Different bodies of law – copyright, patent, privacy, confidentiality and contract law – will be relevant to the collection, storage and dissemination of data. Proper management of data requires an understanding of how these legal regimes impact on the data's generation, handling and dissemination. By adopting mechanisms such as data management policies, plans and toolkits, researchers and research organisations can effectively manage the data they collect or generate, based on a practical understanding of how the various legal regimes apply to it. Implementation of such measures will ensure that research data can be made available online to other researchers in a manner that is openly accessible, timely and in compliance with legal requirements.

OPEN DATA FOR GLOBAL SCIENCE¹

Paul Uhler² and Peter Schröder³

INTRODUCTION

The global science system stands at a critical juncture. On the one hand, it is overwhelmed by a hidden avalanche of ephemeral bits that are central components of modern research and of the emerging ‘cyberinfrastructure’⁴ for *e*-Science.⁵ The rational management and exploitation of this cascade of digital assets offers boundless opportunities for research and applications. On the other hand, the ability to access and use this rising flood of data seems to lag behind,

¹ ‘Open Data for Global Science’ was originally published in *Open Data for Global Science – Special Issue*, Paul Uhler (ed), *CODATA Data Science Journal*, (2007) page 36

<http://www.jstage.jst.go.jp/article/dsj/6/0/OD36/_pdf>. The views expressed in this paper are those of the authors and not necessarily those of their institutions of employment.

² Director of the Office of International S&T Information Programs (ISTIP) and the U.S. National Committee for CODATA at the National Academies in Washington, DC National Research Council.

³ Data Archiving and Networked Services (DANS).

⁴ The US Blue Ribbon Advisory Panel on Cyberinfrastructure anticipated an information and communication technology (ICT) infrastructure of ‘...digital environments that become interactive and functionally complete for research communities in terms of people, data, information, tools and instruments and that operate at unprecedented levels of computational, storage and data transfer capacity...’ in *Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue Ribbon Advisory Panel on Cyberinfrastructure*, National Science Foundation (2003)
<http://www.communitytechnology.org/nsf_ci_report/>. We use the terms cyberinfrastructure and ICT infrastructure interchangeably in this paper.

⁵ ‘*e*-science’ refers to ‘the large-scale science that will increasingly be carried out through distributed global collaborations enabled by the Internet. Typically, a feature of such collaborative scientific enterprises is that they will require access to very large data collections, very large scale computing resources and high performance visualisation back to the individual user scientist . . . Besides information stored in Webpages, scientists will need easy access to remote facilities, to computer – either as dedicated Teraflop computers or cheap collections of PCs – and to information stored in dedicated databases.’ John Taylor, Director General of UK Research Councils. See: <www.research-councils.ac.uk/escience/>.

despite the rapidly growing capabilities of information and communication technologies (ICTs) to make much more effective use of those data. As long as the attention for data policies and data management by researchers, their organisations and their funders does not catch up with the rapidly changing research environment, the research policy and funding entities in many cases will perpetuate the systemic inefficiencies, and the resulting loss or underutilisation of valuable data resources derived from public investments. There is thus an urgent need for rationalised national strategies and more coherent international arrangements for sustainable access to public research data, both to data produced directly by government entities and to data generated in academic and not-for-profit institutions with public funding.

In this chapter, we examine some of the implications of the ‘data driven’ research and possible ways to overcome existing barriers to accessibility of public research data. Our perspective is framed in the context of the predominantly publicly funded global science system. We begin by reviewing the growing role of digital data in research and outlining the roles of stakeholders in the research community in developing data access regimes. We then discuss the hidden costs of closed data systems, the benefits and limitations of openness as the default principle for data access, and the emerging open access models that are beginning to form digitally networked commons. We conclude by examining the rationale and requirements for developing overarching international principles from the top down, as well as flexible, common-use contractual templates from the bottom up, to establish data access regimes founded on a presumption of openness, with the goal of better capturing the benefits from the existing and future scientific data assets. The ‘Principles and Guidelines for Access to Research Data from Public Funding’ from the Organisation for Economic Cooperation and Development (OECD), reported on in another article by Pilat and Fukasaku,⁶ are the most important recent example of the high-level (inter)governmental approach. The common-use licenses promoted by the Science Commons are a leading example of flexible arrangements originating within the community. Finally, we should emphasise that we

⁶ In *Open Data for Global Science – Special Issue*, Paul Uhler (ed), *CODATA Data Science Journal*, (2007).

focus almost exclusively on the policy—the institutional, socioeconomic, and legal aspects of data access—rather than on the technical and management practicalities that are also important, but beyond the scope of this article.

THE GROWING ROLE OF DIGITAL DATA IN THE RESEARCH PROCESS

The evolution of scientific research may be characterised by an accelerating growth in scale, scope, and complexity. These developments in scientific research have been accompanied by a substantial rise in costs. Overall expenditures on research and development (R&D) in the OECD countries increased from \$163.2 billion in 1981 to \$679.8 in 2003 (in constant prices, 2000 dollars: from \$276.6 billion in 1981 to \$638 in 2003).⁷

Not surprisingly, these trends also have elicited growing governmental policy involvement in scientific research at both the national and international levels. The research policy establishment has promoted greater cooperation between public researchers and the private sector, as well as greater international cooperation in public research.⁸ The phenomenal growth of the cyberinfrastructure, particularly in OECD countries, has been both a facilitator and accelerator of these trends. It has further magnified the scale, scope, and complexity of scientific research by enabling the integration of research participants and information resources from multiple disciplines, sectors, and countries.

Continuously growing quantities of data about the universe around us are produced by government agencies, research institutions, and industry as a fundamental component of scientific research worldwide. Practically anything used for research purposes can be described and stored in a digital database. A genomic sequence, the speed of subatomic particles, a response in a social survey, the frequency of nouns in a text corpus, and satellite images of other planets all are used as research data. As described in the National Research Council

⁷ Organisation for Economic Co-operation and Development, *OECD Main Science and Technology Indicators* (2005).

⁸ See, for example, Organisation for Economic Co-operation and Development, *The Knowledge-based Economy* (1996).

symposium on *The Role of Scientific and Technical Data and Information in the Public Domain* in 2002:

The rapid advances in digital technologies and networks over the past two decades have radically altered and improved the ways that data can be produced, disseminated, managed, and used, both in science and in all other spheres of human endeavour. New sensors and experimental instruments produce exponentially increasing amounts and types of raw data. This has created unprecedented opportunities for accelerating research and creating wealth based on the exploitation of data as such ... There are whole areas of science, such as bioinformatics in molecular biology and the observational environmental sciences, that are now primarily data driven. New software tools help to interpret and transform the raw data into unlimited configurations of information and knowledge. And the most important and pervasive research tool of all, the Internet, has collapsed the space and time in which data and information can be shared and made available, leading to entirely new and promising modes of research collaboration and production.⁹

The production of a data set thus constitutes the first stage of improving the knowledge of some part of nature and society for further research and innovation. Rather than a linear process, however, the use of digital data is better conceptualised as a series of dynamic ‘chain link’ feedbacks, broadening the usability of separate and related chains (see Box 1). The increasing supply of data frequently may be useful for purposes beyond those contemplated in the original collection. Many publicly funded data can be of great value for reuse by a broad range of public and private researchers, other types of socioeconomic applications, and the general public.

⁹ Paul Uhler, ‘Discussion Framework’ in Julie Esanu and Paul Uhler (eds), *The Role of Scientific and Technical Data and Information in the Public Domain* (2003) 3.

BOX 1

Research data: their place in the research process

For most of the history of science, scientific data were usually inextricably embedded in an all-embracing research process. Researchers mostly collected and used their own data in their own research projects and had access to few external data sources. However, with the advent of digital technologies and networks, together with the growing scale and scope of research activities worldwide, the various parts of the research trajectory have been loosened into separate specialised activities (as, for example, data collection or technical support) that may be executed by different entities, in-house or outside the research institute. In large-scale research, specialised data service institutes may operate independently from the research projects they serve. Different parties will have differing responsibilities and may have differing claims on ‘their’ parts of the trajectories. The various phases of the research process, including the upstream data management process, may be subject to different policies, regulations, and legislation. This diagram shows the main elements of the research and data trajectories.

The Research Trajectory

The general outline:

Data ⇌ Information ⇌ Knowledge

Detailed stages in the process with feedback (↺, ↻) chains:

hypothesis ↻ (methodology/design) ↻ data collection/measuring ↻ analysis/synthesis ↻ results/publications/patents

The Data Trajectory

Possibilities for data sharing once primary data have been collected:

data collection ⇌ primary data ⇌ processing ⇌ documenting ⇌ final data ⇌ dissemination & archiving
 ↻ **data sharing options**

The changes in the research process have not only been quantitative, but qualitative as well, leading to discoveries never before possible. For example, hitherto unconnected data elements can be assembled into unexpected new results. The research strategy developed by Rita Colwell, former Director of the U.S. National Research Foundation, in

her studies on cholera is a case in point.¹⁰ By combining large sets of data on sea life, earth observation, historical epidemiology, DNA analyses, and social anthropology she was able to demonstrate disease patterns that, without the use of ICT tools and access to all the diverse data, would have remained invisible. What is clear is that digital data play a central part in the emerging global science system and in the promise of *e*-Science. And while most of the palpable progress to date has occurred in the more economically developed countries, the biggest payoffs from this new research paradigm could take place in the developing world.

These major changes in the structure and conduct of data-driven research using the cyberinfrastructure result in an increasing need for rational organisation and planning, however. A more transparent and predictable environment for access to and use of data resources would help to optimise the national and international research system.

THE EMERGING ROLES OF STAKEHOLDERS IN THE GLOBAL SCIENCE SYSTEM IN DEVELOPING DATA ACCESS REGIMES

Changes in the scientific research process are coupled with changing roles of the interdependent parties responsible for science policy and research management. Here we briefly examine the roles of these different stakeholders with regard to public science data policy and management in the context of the cyberinfrastructure. There are formal organisations, associations, and individuals involved at different (inter)national levels in the digital data activities. They represent specific economic, social, national, personal, and scientific interests, and play roles as experts and managers of research. These stakeholder groups all affect the development (or not) of data access regimes, both directly through governmental and institutional data management and policy implementation, and indirectly through normative and behavioural influences.

¹⁰ Rita Colwell, 'A Global Thirst for Safe Water: The Case of Cholera' (Speech delivered at the Abel Wolman Lecture at the National Academy of Sciences, 2002)
<http://www7.nationalacademies.org/wstb/2002_Wolman_Lecture.pdf>.

Governments are responsible for the legal and regulatory framework in which the research system operates, as well as for funding it with the taxpayers' money. Governments have core responsibilities for general public information rights, including overall policy over national science systems. More specifically, governments claim responsibility for overall policy over national science and innovation systems as a public good (for example research for public health, national security, general advancement of knowledge, and socioeconomic development). As funders of research, they have an interest in promoting accountability for the cost effectiveness and management of their public investments in research. Governmental policies are thus crucial for establishing a rational framework for managing and implementing the national science system and international scientific cooperation, most of which is now entirely dependent on digital networks. To the extent that public scientific data (and other types of information) are fundamental components of the modern research enterprise, governments have a responsibility to establish the policy framework in which the research organisations function and enable the rational development and exploitation of those information resources. This involves a balance between protecting and stimulating competitive and cooperative values at different levels of the research system.

Research funding agencies are responsible for the actual allocation of taxpayer funds to the various research activities. They are accountable for the support and performance of the national science system. They comprise the experts who must develop and implement national research strategies and funding priorities in consultation with key representatives of the scientific community. Research funding agencies are also responsible for the more detailed allocation of public research funds, the support of specific elements of the research infrastructure (the people, facilities, and equipment), and the formation of policies specific to their constituencies. Digital science increasingly requires such specific policy and infrastructure support for networks, computing facilities, and institutional mechanisms for storing and making available the digital inputs and outputs of public research. This responsibility includes the possible establishment of specialised data centres, both within the funding agencies themselves and with their support at other research institutions. As the research funding agencies decide on the funding priorities, they are in a powerful position to influence the overall

data policy and management regimes for the research institutions that they create or support.

Universities and not-for profit research institutes manage their employees' implementation of publicly-funded research programs and projects, subject to academic norms and the guidance of the sources of their funding (both public and private, and internal and external). These functions include support and management of ICT facilities and the resulting data collections and repositories for publications. Many academic research institutions now manage a large number of individual databases—as well as specialised data centres and more comprehensive institutional repositories and libraries—that are funded in whole or in part with public money. Whether or not they do have a data centre, they have a responsibility for establishing policies for the access to and use of their expanding amounts and types of research data and information. These policies must be consistent with the requirements and interests of their funding sources, researchers, and other institutional stakeholders, and with the broader research community in which these institutions operate. Widespread uncertainty about possible conflicting interests and tasks of multiple stakeholders make the establishment of data access policies at research institutions crucial, though difficult. They require consistency at the higher policy level, as well as flexibility at the implementation level.

Learned and professional societies represent the formal side of the otherwise more loosely defined research communities. They provide a focal point for interaction and communication by their particular discipline communities, especially at the national level. They are major players in developing scientific norms, values, and standards such as academic freedom, scientific responsibilities, and increasingly regarding access to data produced by members of their research communities. They provide concentrated expert resources that combine the perspectives of the larger-scale changes in the operation of the science system with the first-hand experience from the specific changes in the day-to-day research practice in their discipline areas. The societies promote their views within their own communities by establishing formal and informal policies and codes of conduct for their members, through major conferences and their journal publications, and externally through interactions with policy makers and research managers.

International scientific organisations have a role similar to the learned societies, but at regional or global levels. The international non governmental scientific organisations (NGOs) must be distinguished from the intergovernmental organisations (IGOs). Among the IGOs relevant in this context are the Organisation for Economic Co-operation and Development (OECD), and some of the specialised agencies of the United Nations, such as the United Nations Educational, Scientific, and Cultural Organisation (UNESCO). Relevant NGOs include the International Council for Science (ICSU), the interdisciplinary Committee on Data for Science and Technology (CODATA), the InterAcademy Panel on International Issues (IAP), and the Academy of Sciences for the Developing World (TWAS). These organisations have the subject matter interest and expertise to develop improved data policies and practices, as well as important contacts with the policy and research communities to promote them.

Industry research institutions generally benefit from greater access to scientific data produced by others. Traditionally, industrial laboratories and researchers tend to keep their own data outputs proprietary and inaccessible to other scientists and engineers. Keeping proprietary data inaccessible might entail lost opportunity costs for the owners as they will not be able to benefit from the results of additional research by other experts using those data. Industry research institutions increasingly outsource research to universities, however, partnering with university researchers often keeping the data on a proprietary basis. Industry-academic research partnerships are growing because of public policies favouring such arrangements and economic pressures on both academic and industrial research organisations. Public-private research partnerships may further complicate the management of the resulting data and the optimal allocation of rights to those data, as discussed further in the article.

Individual researchers generate increasing amounts and types of data, both as individuals and as participants in various kinds of formal and informal collaborations. Individual researchers sometimes show a different attitude to accessing data from colleagues for their own research than towards sharing 'their' data with colleagues. The informal culture at the working research level, with its strategic relations among researchers that are often invisible to outsiders, is dominated by traditions that in many cases have not yet caught up with recent

developments in data policies and data management. However, much of the formal decision making on data access and sharing increasingly takes place at the institutional level. As the main producers and users of public scientific data, individual researchers ultimately have the greatest stake in the development of rational data access regimes and in the adequate funding and management of data collections and centres. Because researchers typically have been at the forefront of both developing and using the ICT infrastructure, they also have been some of the most influential players, together with their employing institutions, in creating new models of data access regimes from the bottom up. A great deal of data exchange and collaboration takes place informally on the internet between scientists as a result of their personal and professional relationships and in support of their respective research activities. Many researchers also have become part-time or specialised data managers.

The general public includes the taxpayers whose money is invested in public research and related data activities. Society in general has a strong interest in seeing that the fruits of those investments are effectively managed and used. The lay public generally is not concerned directly with the policy and management issues pertaining to national R&D, or to data from publicly funded research. Nevertheless, action groups of citizens may get involved in data access issues for various specific reasons and circumstances (e.g., local environment, health, or consumer safety). Increasingly, journalists do their own analyses of datasets used in the social sciences and the humanities. Moreover, with the broad public access to the internet in many countries, the potential user base for many kinds of public research data has expanded greatly, adding a further important dimension to the data policy debate, as discussed further on in the article.

Each of these major stakeholder groups in the research enterprise has a major and growing interest in the development of more effective policies for access to and use of publicly funded research data. Although the sharing of data resources in networked cooperation has become standard practice in some fields, particularly in the more economically developed countries, in many cases researchers and their institutes experience too much uncertainty and barriers to make the most effective use of the new possibilities. This situation is exacerbated in less developed countries that also have less fully developed technical and

human infrastructure for research, as well as institutional mechanisms and policy frameworks.

THE HIDDEN COSTS OF CLOSED DATA SYSTEMS

As described in Box 1, digital research data are emerging in the research system as autonomous resources, the uses of which are no longer tied solely to their original producers or purposes. There are, of course, data that have little value outside the narrow research project for which they were collected or that are not useful for lack of quality, insufficient documentation, or other deficiencies. Many types of data, however, can be used beyond the ambit of the original producers and users in diverse and unlimited ways, at different times and places, and potentially by anyone with access to the ICT infrastructure. The sharing of public research data opens up new opportunities to raise the quality and productivity of research, but the full realisation of this potential requires additional attention to data policy and practice.

At the same time, there are competitive values and other legitimate reasons for restricting access to data from publicly funded research, which is reviewed further on in the article. The different stakeholders involved may perceive conflicting interests when considering the benefits and drawbacks of open access to data. Many researchers tend to treat the data they produce through publicly-funded research as individual or institutional property, and this view frequently is reinforced by the lack of adequate policy guidance from their public funding sources.

There are, however, a number of negative implications¹¹ to the efficiency and effectiveness of the research system from unnecessarily balkanised and closed access regimes in light of the (quasi) public good¹² nature of such digital data resources.

¹¹ J Reichman and Paul Uhler, 'Database Protection at the Crossroads: Recent Developments and Their Impact on Science and Technology' (Spring 1999) 14 (2) *Berkeley Technology Law Journal* 819–21.

¹² Both the public nature of the research and the resulting data have public-good characteristics. A public good is both non-rival and non-excludable. The former means that it costs nothing to provide the good to another person once someone has produced it (in other words, it has a zero marginal cost of distribution). The latter refers to the characteristic that once such a good is produced, the producer cannot exclude others from benefiting from it. Inge Kaul et al, 'Defining Global Public Goods' in Kaul et al (eds), *Global Public Goods: International Cooperation in*

Higher research costs

Most obviously, restricting access imposes structural inefficiencies and higher research costs. Many factual databases cannot or should not be independently recreated, either because they contain observations of unique phenomena, historical information, or cost a great deal to generate.¹³ Moreover, databases with a monopoly status that are maintained on a closed proprietary basis will tend to result in higher, anti-competitive pricing.¹⁴ Managing publicly funded databases on a restrictive, proprietary basis also adds substantial administrative overhead on both ends to make each transaction, further taxing the public research system. This is particularly exacerbated by public institutions that license data at high costs and restrictions to other public institutions.

Lost opportunity costs

Perhaps not as obvious, there is much less data-intensive research possible if the publicly-funded data are not shared or made easily available online. This results in significant lost opportunity costs that are certain to occur, but are difficult to measure.¹⁵ A simple analogy might suffice to illustrate this effect. Just as it would hardly be cost-effective research management to limit the use of a telescope or an accelerator to the researchers and engineers who designed the instrument, it is a waste of effort and money to limit the use of data to the researchers

the 21st Century (1999). Public research and publicly funded scientific data on digital networks may be considered as 'quasi public goods' in that they are to a certain degree appropriable, although they nonetheless have public-interest characteristics that make them capable of production only if subsidised by public funding. See Michael Callon, 'Is Science a Public Good?' (1994) 19 *Science, Technology and Human Values* 395.

¹³ National Research Council, *A Question of Balance: Private Rights and the Public Interest in Scientific Databases* (1999) 19–20.

¹⁴ Peter Weiss, 'Conflicting International Public Sector Information Policies and Their Effects on the Public Domain and the Economy' in Julie Esanu and Paul Uhler (eds), *The Role of Scientific and Technical Data and Information in the Public Domain* (2003) 129–32; and J Reichman and Paul Uhler, 'Database Protection at the Crossroads: Recent Developments and Their Impact on Science and Technology' (Spring 1999) 14 (2) *Berkeley Technology Law Journal* 819–21.

¹⁵ It is difficult to determine what might have been possible if only the data were openly available. This was analysed in at least one instance when the U.S. Landsat program was privatised in the mid-1980s. National Research Council, *Bits of Power: Issues in Global Access to Scientific Data* (1997) 121–24.

responsible for their original collection and lose the potential benefits of greatly expanded applications for those data that may have some broader utility.

Barriers to innovation

The production downstream of copyrightable or patentable intellectual goods by both the public and private sectors depends to a large extent on access to the free flow of upstream public factual data and information. The overprotection or unavailability of public databases leads to deadweight social costs, taxing the innovation system in each country and slowing scientific progress.¹⁶

Less effective cooperation, education, and training

A failure to make research data easily available, or erecting barriers that are too high, necessarily results in less effective interdisciplinary, inter-institutional, inter-sectoral, and international cooperation. In the same way, students may be less effectively educated and trained if they are unable to work with a broad cross-section of data. These barriers are reinforced in many cases by myopic policies that provide access and restricted use for a small number of pre-approved investigators formally associated with specific research projects and programs, even at an international level, while greatly constraining both access and use of those data by researchers and other potential users in ‘non-approved’ disciplines, institutions, sectors, and nations.

Sub-optimal quality of data

Data organised in a closed environment frequently will be subject to a process of validation and verification from a substantially smaller and less diverse scientific community than data that are openly available. This will increase the risks of lower data quality and consequently of the quality of research outcomes. Less comprehensive opportunities for quality control will diminish the return on investments in data as well as research.

¹⁶ J Reichman and Paul Uhler, ‘A Contractually Reconstructed Research Commons for Scientific Data in a Highly Protectionist Intellectual Property Environment’ (Winter/Spring 2003) 66 *Law and Contemporary Problems* – *Duke University School of Law* 410–16.

Widening gap between OECD nations and developing countries

Developing countries are particularly disadvantaged by a lack of availability or high barriers to access. Although not all databases produced in OECD countries are relevant in less developed ones, either because of their subject matter or geographic focus, those that do have broad applicability as a global public good will typically be unused in the developing world if there is a high price for access, and in many cases, any charge at all.

Unnecessary access barriers to publicly funded research data therefore result in diminished returns on the social and scientific capital investments in public research and in the inefficient distribution of benefits from those investments, even as the improving technological capabilities offer ever greater opportunities to increase that return.

THE SCIENTIFIC AND SOCIOECONOMIC BENEFITS OF GREATER OPENNESS

In view of the trends and the role of public data in science discussed above and the inefficiencies of the current ad hoc system, there are many compelling reasons for developing more comprehensive access regimes at the institutional, national, and international levels, with open access as the default rule. This is the case whether the data are produced within government or by entities funded by government sources, although some important distinctions apply, as outlined below.

Open access in the context of public research data may be defined as access on equal terms for the international research community, as well as industry, with the fewest restrictions on (re)use, and at the lowest possible cost.¹⁷

This definition is also consistent with the ‘full and open’ data policy used in various international environmental projects and in environmental (and other) research in the United States over the past two decades.¹⁸

¹⁷ Preferably at no more than the marginal cost of dissemination (the cost of fulfilling a user request), which is essentially zero online.

¹⁸ National Research Council, *Bits of Power: Issues in Global Access to Scientific Data* (1997) 1, 15–16.

Because the value of scientific data lies in their use, open access to and sharing of data from publicly-funded research offers many advantages over a closed, proprietary system that places high barriers to both access and subsequent re-use. Open access to such data:

- reinforces open scientific inquiry,
- encourages diversity of analysis and opinion,
- promotes new research and new types of research,
- enables the application of automated knowledge discovery tools online,
- allows the verification of previous results,
- makes possible the testing of new or alternative hypotheses and methods of analysis,
- supports studies on data collection methods and measurement,
- facilitates the education of new researchers,
- enables the exploration of topics not envisioned by the initial investigators,
- permits the creation of new data sets, information, and knowledge when data from multiple sources are combined,
- helps transfer factual information to and promote capacity building in developing countries,
- promotes interdisciplinary, inter-sectoral, inter-institutional, and international research, and
- generally helps to maximise the research potential of new digital technologies and networks, thereby providing greater returns from the public investment in research.¹⁹

Open access to factual data plays a vital enabling role in all these areas. Creating a level playing field for researchers and their institutes is

¹⁹ See, for example, S E Feinberg, M E Martin, and M L Straf (eds), *Sharing Research Data* (1985); National Research Council, *A Question of Balance: Private Rights and the Public Interest in Scientific Databases* (1999); and Arzberger et al, 'Promoting Access to Public Research Data for Science, Economic, and Social Development' (2004) *CODATA Data Science Journal*, 135–52.

impossible without broad and effective access to publicly funded research data. Nevertheless, there are essential distinctions to be made between data produced by government entities and by entities funded by government sources, as well as across disciplines and types of data. Moreover, there may be important and legitimate reasons for not making publicly funded research data openly accessible, but rather keeping them secret or proprietary, at least for limited times and in specific circumstances. These nuances and exceptions are complex, but important to understand in the development of access regimes. We touch on them only briefly below.

Policy Considerations for Data Produced by Government Entities

The data and databases generated directly through government research have the following additional policy considerations favouring their open availability and unrestricted reuse²⁰:

Legal considerations

Consistent with Article 19 from the *Universal Declaration of Human Rights*, national law on information rights should include public access to data and information produced by the government, and related freedom of expression by the public. Moreover, a government entity needs no legal incentives from exclusive property rights to create the data. Both the activities that the government undertakes and the information produced by it in the course of those activities are a public good, properly in the public domain. Data produced through public research frequently have global public-good characteristics.²¹

Socio-economic considerations

Open access is the most efficient way to disseminate public data and information online in order to maximise the value and return on the

²⁰ Paul Uhler and UNESCO, *Policy Guidelines for the Development and Promotion of Governmental Public-Domain Information* (2004) 49.

²¹ See, for example, Dana Dalrymple, 'Scientific Knowledge as a Global Public Good: Contributions to Innovation and the Economy' in Julie Esanu and Paul Uhler (eds), *The Role of Scientific and Technical Data and Information in the Public Domain* (2003) 35–51.

public investment in its production.²² There are numerous economic and non-economic positive externalities—especially through network effects—that can be realised on an exponential basis (though they may be difficult to quantify) through the open dissemination of public-domain data and information on the internet.²³ Conversely, the commercialisation of public data on an exclusive basis produces de facto public monopolies that have inherent economic inefficiencies and tend to be contrary to the public interest.

Ethical considerations

The public has already paid for the production of the information. The burden of fees for access falls disproportionately on the poorest and most disadvantaged individuals (and researchers), including those in developing countries when the information is made available online. This is an important consideration for public, governmental scientific data that constitute a global public good.

Good governance considerations

Transparency of governance is undermined by restricting citizens from access to and use of public data and information created at their expense and on their behalf. Rights of freedom of expression are compromised by restrictions on re-use and re-dissemination of public information. It is no coincidence that the most repressive political systems make the least amount of government information, especially factual data, publicly available.

Although there are strong arguments in favour of a default rule of openness in support of publicly-funded research, at the same time there are various legitimate, countervailing policies that may limit the free and unrestricted access to and use of government information, including

²² Joseph Stiglitz et al (commissioned by the Computer and Communications Industry Association), *The Role of Government in a Digital Age* (2000).

²³ Joseph Stiglitz et al (commissioned by the Computer and Communications Industry Association), *The Role of Government in a Digital Age* (2000). See also Peter Weiss, 'Conflicting International Public Sector Information Policies and Their Effects on the Public Domain and the Economy' in Julie Esanu and Paul Uhler (eds), *The Role of Scientific and Technical Data and Information in the Public Domain* (2003) 129–32; Commission of the European Communities (European Union), *Public sector information: A key resource for Europe* (1998); and PIRA International for the Directorate General for the Information Society (European Union), *Commercial Exploitation of Europe's Public Sector Information*, Final Report (2000).

research data. For example, there are statutory exemptions to public access and use based on national security and law enforcement concerns, the need to protect personal privacy, and to respect confidential information (plus other exemptions to Freedom of Information laws, where applicable).²⁴ Government agencies also should respect the proprietary rights in information originating from the private sector that are made available for government use, unless expressly exempted. Governments may adopt policies as well against competing directly with the private sector in providing certain information products and services. 'Emerging Open Access Models' examines more explicitly some of the additional factors that need to be considered in limiting disclosure of data in research funded by the government.

Policy Factors to Consider in Disseminating Government-Funded Research Data

The access policies for research data produced by non-governmental entities with government funds²⁵ have rationales similar to those outlined above for government-produced data. There are additional factors that may come into play, however.

In some areas of research or in certain research programs, the recipient of a government grant or contract may have a specifically established period of exclusive use of the research data or until publication of the research results. These policies vary across disciplines, institutions, and countries, and in many cases there are no expressly stated, formal rules, just community practice and norms. In some instances, it is appropriate for data to be withheld even after publication, either because of confidentiality or privacy requirements, or because the underlying data are part of a longitudinal study spanning many years. However, generally accepted scientific norms and the exigencies of the scientific process that require access to data underlying published results for the purpose of independent verification make disclosure of such data

²⁴ For a compendium of freedom of information laws and their exceptions, see <<http://www.freedominfo.org>>.

²⁵ This is certainly the case in which public sources provide 100 percent of the funding. As the percentage of public funding in any given research project diminishes the corresponding rationale and arguments for full policy control become weaker as well.

following publication an essential prerequisite for sound science, even if there is no formal rule in place.²⁶

Moreover, open access to research data will not in itself result in usability. Optimum accessibility and usability presuppose a trajectory of proper organisation and curation of a database with ‘added’ value, which also adds costs to its production. Investments in preparing factual data for broader use may easily qualify for intellectual property protection and require some source of funding for providing enhanced access to other users. In most cases, however, there is a compelling reason to develop legal and funding mechanisms that will actively promote public accessibility to those publicly funded data resources. Such complications strengthen the case for further cooperation among the different parties involved in developing the policies and institutional mechanisms for improved data management and access.

Some OECD countries or research funding agencies also have policies that favour the commercialisation of government-funded research.²⁷ For research areas in which commercial applications are inherent or desirable, there will be additional motivations for the researcher to keep the data proprietary and under conditions of trade secrecy, at least until patent rights are secured. Furthermore, the non-governmental research may involve a mix of public and private funds or partners, or include parties from multiple countries, which can complicate the allocation of rights in the research data. In such cases, the application of an open access data policy also may be inappropriate, unless expressly agreed to by all the participating parties.

The issues raised in public-private relationships take many forms and contain some inherent tensions, such as openness versus exclusivity, public goods versus private investments, public domain versus proprietary rights, and competition versus monopoly, among others. This mix of motivations, priorities, and requirements is context-

²⁶ See, for example, National Research Council, *Community Standards for Sharing Publication-Related Data and Materials* (2002).

²⁷ Perhaps the best known of these is the 1980 ‘Bayh-Dole Act’ in the United States, which states in part: ‘[i]t is the policy and objective of Congress to use the patent system to promote the utilisation of inventions arising from federally supported research or development...[and] to promote the collaboration between commercial concerns and non-profit organisations, including universities...’, Public Law No 96–517, § 6(a), 94 Stat 3015 (1980), codified as amended at 35 USC, § 200.

dependent, typically unique to the parties involved, and frequently not amenable to inflexible statutory and regulatory frameworks. In such cases, the ordering of the respective rights and interests of the parties involved is most efficiently accomplished through contracts. Such private agreements provide maximum flexibility within the larger research policy context. What is especially important to emphasise here is that such agreements can in many cases provide for conditionally open access that advances the public interest goals associated with the public funding, while effectively protecting existing proprietary private interests.²⁸

This bifurcated ordering of interests can take many forms. At the most basic level, it is possible to provide free access for not-for-profit research and education (and other) users, while restricting commercial users and uses to a reimbursable, or even for-profit, basis. Various techniques of price discrimination and product differentiation may be similarly employed, based on factors such as time (for example, real-time access for commercial users vs. delayed access for non-profits), scope of coverage (for example, geographic or subject matter limitations), levels of customer support or service, and other possible distinctions.²⁹ Such strategies can help promote scientifically and socially beneficial access and use, not only in the complex public-private research relationships, but even in exclusively private-sector settings.³⁰

In addition to these complexities within the government-funded academic and not-for-profit research context, there are important distinctions that need to be made among different disciplines and types of research. A major difference is between those areas of science that are dominated by 'big science' research projects and programs, and those that remain predominately 'small science' research endeavours, performed by a single investigator (or small group).³¹ The former are

²⁸ J Reichman and Paul Uhler, 'A Contractually Reconstructed Research Commons for Scientific Data in a Highly Protectionist Intellectual Property Environment' (Winter/Spring 2003) 66 *Law and Contemporary Problems* – *Duke University School of Law* 410–16.

²⁹ National Research Council, *Bits of Power: Issues in Global Access to Scientific Data* (1997) 124–6.

³⁰ See generally, J Reichman and Paul Uhler, 'A Contractually Reconstructed Research Commons for Scientific Data in a Highly Protectionist Intellectual Property Environment' (Winter/Spring 2003) 66 *Law and Contemporary Problems* – *Duke University School of Law* Part IV.

³¹ Traditionally, 'small science' research was done primarily in experimental laboratory sciences, such as chemistry and biology; in fieldwork studies such as ecology, anthropology, and various

typically cooperative, whereas the latter tend to be more competitive, or at least insular. Most big science programs have instituted a formal data access regime in established data centres, frequently on an open access basis (as discussed further in Emerging Open Access Models), whereas the latter generally have no formal access rules governing their research data.

Another key distinction across scientific disciplines is between the observational and experimental sciences, where the types of data that need to be preserved and made broadly available differ significantly.³² Typically, for observational data sets, it is the raw or minimally processed data that have the greatest value for reuse in research, whereas in the experimental sciences, it is the highly evaluated and verified data that are preserved and made available for broad use.

Finally, as already noted for government-produced data, an important distinction must be made between data collected on human subjects and data on other, impersonal, subjects.³³ Research data on human subjects are restricted in various ways on ethical and legal grounds to protect personal privacy.

The bottom line in all of these categories of research and data types, however, is that open access to publicly funded research data should be the default rule and operating presumption, rather than the exception, and the exceptions to openness should be based on explicit, well-justified grounds.

areas of social science; and in studies of human subjects, such as the biomedical and behavioural sciences. The autonomous nature of the research, and in many cases the privacy concerns associated with human studies, have precluded the sharing of data or the pooling of small data sets in centralised repositories. Here the research has been more competitive than cooperative and any exchanges of data were typically done on an informal, collegial basis, rather than through some formally structured data access regime. With the advent of higher capacity computing and digital networks, however, some of these research areas have organised 'big science' research programs (for example, the human genome project) and become much more data-intensive. They have established their own specialised data centres (for example, genomic and protein data in molecular biology) or formed distributed data networks with nodes (for example, ecological or biodiversity data). J. Reichman and Paul Uhler, 'A Contractually Reconstructed Research Commons for Scientific Data in a Highly Protectionist Intellectual Property Environment' (Winter/Spring 2003) 66 *Law and Contemporary Problems* – *Duke University School of Law* 343–4 and 426–7.

³² National Research Council, *Preserving Scientific Data on Our Physical Universe* (1995) 34–6.

³³ Organisation for Economic Co-operation and Development, *OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data* (1980).

EMERGING OPEN ACCESS MODELS

The presumption of openness and the implementation of an open access policy as the default rule in publicly funded research is certainly not a revolutionary concept. Not only are there solid justifications for such a policy as outlined above, but there are innumerable examples of successful implementations of this policy in practice in both government and government-funded institutions, in many fields of research, and in many countries. In this section we characterise these examples broadly and provide a number of specific references. Box 2 identifies a range of distributed, open, collaborative research and information production and dissemination activities using digital networks,³⁴ while Box 3 provides details about one compelling example, identified in Box 2, of open access to academic materials at a world-class university.

There are many new kinds of distributed, open collaborative research and information production and dissemination on digital networks. Examples of open data and information production activities include:

Box 2

- Open-source software movement (such as, Linux and 10Ks of other programs worldwide, many of which originated in academia and are developed for research purposes);
- Distributed Grid computing or e-Science (such as, SETI@Home, LHC@home);
- Community-based open peer review (such as, Journal of Atmospheric Chemistry and Physics); and
- Collaborative research Web sites and portals (such as, NASA Clickworkers, Wikipedia, Curriki).

The following are examples of open data and information dissemination and permanent retention:

³⁴ Paul Uhler, 'The emerging role of open repositories for the scientific literature as a fundamental component of the public research infrastructure' in G Sica (ed), *Open Access: Open Problems* (2006).

- Open data centres and archives (such as, GenBank, the Protein Data Bank, The SNP Consortium, Digital Sky Survey); Federated open data networks (such as, World Data Centers, Global Biodiversity Information Facility; NASA Distributed Active Archive Centers);
- Virtual observatories (such as, the International Virtual Observatory for astronomy, Digital Earth);
- Open access journals (such as, BioMed Central, Public Library of Science, + > 2500 scholarly journals);
- Open institutional repositories for that institution's scholarly works (such as, the Indian Institute for Science, plus hundreds globally);
- Open institutional repositories for publications in a specific subject area (such as, PubMedCentral, the physics arXiv);
- Free university curricula online (such as, the MIT OpenCourseWare); and
- Emerging discipline-based commons (such as, the Conservation Commons, the Geoscience Information Commons)

*Box 3****The OpenCourseWare initiative at the Massachusetts Institute of Technology***

The digital revolution is transforming information economics in a radical way. In the public science system one of the interesting trends is the development of additional user bases for 'secondary' use of data, information, and knowledge. When openly available, publicly funded digital resources can have many new useful 'lives' in addition to their primary uses. Use of the internet has minimised distribution costs. Open access is a way of cutting transaction costs. Low access barriers serve the original purposes of the public investment and increase the return on the investment: a broader scientific workforce can be put to work to get additional results without investments in additional resources.

Low access barriers make it possible to meet an important demand that cannot be served through traditional markets. For example, in 1999 the Massachusetts Institute of Technology (MIT) investigated a business model for selling its curriculum materials online. When it appeared that there would be an insufficient market for this service, MIT did not abandon the idea, but changed the original business model into one of open access: the 'OpenCourseWare' initiative. The university now offers free access to well over one thousand courses and has gotten hundreds of million hits on its portal from educators, students, and self-learners from all over the world. Of course, the project initially was greeted with a great deal of apprehension among the MIT faculty, but eventually this bold vision was accepted. As expressed by President Emeritus of MIT Charles M Vest: *'OpenCourseWare looks counterintuitive in a market-driven world. But it really is consistent with what I believe is the best about MIT. It is innovative. It expresses our belief in the way education can be advanced – by constantly widening access to information and by inspiring others to participate.'*

Together, these various open access activities constitute an emerging globally networked 'commons' for public science, representing a broad range of information types, institutional structures, disciplines, and countries. A common policy aspect of all these activities is their provision of free and open access online, with either reduced retention

of intellectual property rights through permissive licensing mechanisms³⁵ or, much less frequently, a statutory public domain status.³⁶

In the area of data from publicly funded research, there already are many open access activities throughout the world, although no comprehensive compendium currently exists. As indicated in Box 2 there are at least two major types of institutional models specific to data: (1) open data centres or archives, and (2) federated³⁷ open data networks. The former is a centralised model whereas the latter has a connected set of distributed nodes. There are numerous examples of each type of open

³⁵ For a selection of such permissive licensing templates, which use statutory intellectual property protection, but with only ‘some rights reserved’ instead of all the rights accorded under the statute, see the Creative Commons and its more recent Science Commons initiative <<http://www.creativecommons.org>>.

³⁶ The public domain status of factual data is a complex legal subject. Some countries expressly exclude government-generated information from copyright. Moreover, under traditional copyright law, factual compilations that lacked creativity or originality in their selection or arrangement, like many of the databases that are the subject of discussion in this paper, were not copyrightable and all the data in those compilations were in the public domain. However, some jurisdictions had so-called ‘sweat-of-the-brow’ common-law protections (for example, the United Kingdom and certain states in the United States), while others adopted more formal statutory protection of non-copyrightable compilations (for example, the Scandinavian Catalogue Rule). More recently, the European Union enacted exclusive property protection of databases and compilations of information (*Directive 96/9/EC of the European Parliament and the Council of 11 March 1996 on the Legal Protection of Databases* [1996] OJ L 077), which has been implemented in all E.U. Member States and Affiliated States, as well as in some other countries. This protection in most countries applies even to government and government-funded databases. In most countries there are very limited exceptions for public-interest uses of data (for example, for public scientific research or education), and in some jurisdictions (for example, France, Italy, Greece) there are no exceptions at all. For a comprehensive description and analysis of the E.U. Database Directive and its potential long-term effects of public research, see J. Reichman and Paul Uhler, ‘Database Protection at the Crossroads: Recent Developments and Their Impact on Science and Technology’ (Spring 1999) 14 (2) *Berkeley Technology Law Journal* 819–21; and J. Reichman and Paul Uhler, ‘A Contractually Reconstructed Research Commons for Scientific Data in a Highly Protectionist Intellectual Property Environment’ (Winter/Spring 2003) 66 *Law and Contemporary Problems – Duke University School of Law* 410–16.

³⁷ This type of management structure for distributed scientific data archives and data centres was first described in National Research Council, *Preserving Scientific Data on Our Physical Universe* (1995) 51–3. This model was based on a ‘flat’ corporate management model described in Charles Handy, ‘Balancing Corporate Power: A New Federalist Paper’ (1992) 70(6) *Harvard Business Review* 59–72. The key elements of a federated management model are: subsidiarity (the power is assumed to lie within the subordinate units of the organisation), pluralism (interdependence of members), standardisation of key elements to facilitate cooperation and interoperability, a separation of powers (responsibilities), and strong leadership from a small central directorate that is effective but not overbearing.

access data model operated either directly by government agencies or by government-funded entities (universities and not-for-profit research institutes).

Despite the successful adoption of open data access policies and practices in many areas of public research, the application of such regimes remains fragmented and inconsistent—a patchwork of uncoordinated and largely disparate activities, many of which are ad hoc, bottom-up endeavours. In too many cases, establishing satisfactory arrangements for data access seems to go beyond the means and imagination available at the working level. If finding adequate solutions without outside help is too much trouble, the researchers involved may easily succumb to passive risk avoidance. In view of the potential benefits that can be derived from increasing and improving access to such resources, establishing a more transparent and predictable environment that is coordinated at the national and international levels is desirable.

Some science policy leaders have begun to address these exigencies at the national level. For example, China established the Scientific Data Sharing Program in 2002.³⁸ Canada launched a National Consultation on Access to Scientific Research Data in 2004³⁹ and, that same year, the Research Council of Norway released a white paper documenting the important role of databases as a research infrastructure component.⁴⁰ In 2005, the U.S. National Science Board called for an initiative to develop a national policy framework for long-lived data collections,⁴¹ which was followed up by the establishment of an Interagency Working Group on Digital Data in the White House Office of Science and Technology

³⁸ Jinpei Cheng, 'Development of China's Scientific Data Sharing Policy' in Julie Esanu and Paul Uhler (eds), *Strategies for Preservation of and Open Access to Scientific Data in China* (2006). Also discussed in the article by Guan-hua Xu in *Open Data for Global Science – Special Issue*, Paul Uhler (ed), *CODATA Data Science Journal*, (2007).

³⁹ David Strong, and Peter Leach (National Research Council), *National Consultation on Access to Scientific Research Data* (2005) 82. Also discussed in the article by Sabourin and Dumouchel in *Open Data for Global Science – Special Issue*, Paul Uhler (ed), *CODATA Data Science Journal*, (2007).

⁴⁰ The Research Council of Norway, *The Need for Scientific Equipment, Databases, Collections of Scientific Material, and Other Infrastructure* (2004) report submitted as input to the *White Paper on Research* (2005) Oslo (Abridged English version).

⁴¹ National Science Board (National Science Foundation), *Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century* (2005) 64.

Policy.⁴² Most research funding agencies in the United States also have developed data policy guidelines for their grantees that encourage data sharing or deposits in established community data repositories, within specific discipline or research program contexts. However, the existing institutional policies still remain ad hoc and sub-optimally coordinated at the national level in the United States, as in most other countries.

At the international level, initiatives such as the Budapest Open Access Initiative, the Bethesda Declaration, and the Berlin Declaration,⁴³ although focused more on open access to the scholarly journal literature than to the data, have helped to pave the way for further national policies. The new ‘Guidelines for Access to Research Data from Public Funding’ from the OECD, endorsed by the governments of OECD countries (as discussed towards the end of this paper⁴⁴), may be expected to play an important catalytic role.

While these incipient institutional models and policy approaches are commendable indicators that the scientific community is awakening to the opportunities and challenges of comprehensively rationalised data access regimes in public science, a great deal more can and should be done. And although the patchwork quilt of bottom-up data access regimes has served some research communities well in some cases, this loosely decentralised aggregation of approaches could achieve much greater results from a concerted national and international policy and funding focus.

TOWARD OPEN DATA REGIMES: GUIDING PRINCIPLES AND FLEXIBLE CONTRACTUAL TEMPLATES

The foregoing discussion has sought to develop a rationale for more formalised data access policies and procedures in public research, based

⁴² Declan Butler, ‘Agencies join forces to share data’ (2007) 446 *Nature* 354.

⁴³ The *Budapest Open Access Initiative* (2002) is available at: <<http://www.soros.org/openaccess/read.shtml>>; the *Bethesda Statement on Open Access Publishing* (2003) is available at: <<http://www.earlham.edu/~peters/fos/bethesda.htm>>; and the *Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities* (2003) is available at: <<http://www.zim.mpg.de/openaccess-berlin/berlindeclaration.html>>.

⁴⁴ This is also discussed in an article in *Open Data for Global Science – Special Issue*, Paul Uhler (ed), CODATA Data Science Journal, (2007).

on a core default principle of openness. The benign neglect of research data and databases thus far has not been regarded as a significant policy blunder. The most pressing database requirements seem to have been met through the ad hoc resourcefulness and volunteerism of dedicated individuals in public science.⁴⁵ But the brief history of the digital age already is replete with major losses of data and missed opportunities⁴⁶ that are certain to multiply in the absence of sustained focus and action.

As previously discussed, it also is important to recognise that public policies in the developed and developing countries alike are shaped by legitimate considerations and interests that do not leave all scientific information and data in the public domain or under pure open access conditions. Instead, they impose limitations upon openness and cooperation in the conduct of public research and the utilisation of its findings, in varying degrees and for a variety of purposes. Consequently, there is a need for public policies and institutional arrangements to seek a judicious balance between positive and negative effects upon the conduct of publicly funded research that are likely to ensue from the granting and enforcing of private ownership rights in scientific and technical data and information. Yet, in recent decades the policy balance in this regard has been disrupted in ways that some science policy analysts perceive as threatening the long-term vitality of fundamental scientific research.⁴⁷

A successful data access regime must involve a comprehensive framework of policies and procedures that are based on a complete set of supporting principles and guidelines. Areas that require attention in developing principles and subsequent access regimes include organisational and management, financial and economic, legal, socio-cultural, and technical considerations.⁴⁸ The costs of inaction in the current state of affairs continue to accumulate, while the opportunities

⁴⁵ Stephen Maurer, Richard Firestone and Charles Sriver, 'Science's neglected legacy' (2000) 405 *Nature*.

⁴⁶ See, for example, National Research Council, *Bits of Power: Issues in Global Access to Scientific Data* (1997) 121–4.

⁴⁷ See, for example, Paul Uhler, 'Discussion Framework' in Julie Esanu and Paul Uhler (eds), *The Role of Scientific and Technical Data and Information in the Public Domain* (2003) 129–32.

⁴⁸ Arzberger et al, 'Science and Government: An International Framework to Promote Access to Data' 303 *Science* 1777–8.

provided by the emerging cyberinfrastructure and new science initiatives will remain suboptimal.

Because of the diverse role of data in different fields of research, and the diverse and sometimes competing interests of the different stakeholders in the research enterprise, the formal data regimes need to be tailored to specific circumstances, but managed for the greatest return on the public investments. These conditions make it essential for most policy directives from the top at the national and international levels to be flexible and not rigidly prescriptive, while providing sufficiently strong and comprehensive guidance to the entities at the working level to implement effective regimes that are responsive to their particular interests.

In this final section we examine some mechanisms that can improve top-down guidance on the one hand, and bottom-up flexibility on the other. The former are the high-level international principles that can help guide the development of specific data access regimes at the (inter)national level. The latter involve the practical implementation through the development and voluntary adoption of new licensing templates that rights holders can select as standard options to provide access and use on less restrictive terms and conditions. We conclude with a brief overview of a major new initiative that seeks to integrate more effectively the top down and bottom up approaches.

Guiding principles

A good starting point for regulation at the more general level is the development of international principles, based on consensus by the national participants, which can help provide guidance to the governments, the public agencies, institutions, and individual researchers engaged in publicly funded research worldwide.⁴⁹ Coherent, consensus-

⁴⁹ One example of this type of consensus-building international process is the OECD *Ministerial Declaration on Access to Research Data from Public Funding* of 30 January 2004 and the 2007 *OECD Guidelines* that followed it, as described by Pilat and Fukasaku in *Open Data for Global Science – Special Issue*, Paul Uhler (ed), *CODATA Data Science Journal*, (2007). The *Declaration* was inspired by the successful examples of data sharing on the (inter)national and institutional levels. The science ministers agreed that OECD guidelines would contribute to reach common science policy goals by improving the quality and productivity of scientific research and increasing the cost effectiveness of public investment in scientific research. The essence of the *Declaration* lies in the Principles that systematically treat the main points of the data access issues that have been worked out in subsequent *Guidelines*.

based international principles, building on the experience of established successful models, should provide a number of benefits. They indicate the collective importance placed by science leaders in the national governments to the public research data issues. They can articulate a rationale and responsibility for improving the management and funding of the public data resources. They can provide guidance for the development of new access regimes based on a common set of values and objectives. And they can help establish an international level playing field for research and industry. The end result may be expected to lead to a higher return on public investments in research and substantial increases in productivity and cost-effectiveness.

The development of overarching international principles that cover publicly-funded research data in many countries can only be restricted to the essentials, of course. In the many different countries, disciplines, and institutes complete compliance with the principal rules will be difficult, and there will always be exceptions to the rules. Context-dependent solutions will have to be found, but all of these exceptions cannot and should not be part of the principles. The perspective can only be that of stating the *default* rules, including the core openness principle. Applying the principles and working out the specific details will be the responsibility of the stakeholders identified above—the national governments, public research funding agencies, and universities and public research institutes—in collaboration with the research community, as represented by the learned societies and the private sector. The principles therefore should offer the general international guidance for further regulation by the parties more directly involved.

The principles should not conflict with national legislation, nor harm other national, institutional, or individual interests. Strong, simple principles should be distilled from a much more extensive body of input and from a broad consultative process.

At the level of international science policy, principles represent the broadest common denominator of existing policies and (best) practices. But from this common ground they should guide emerging processes of change. International principles ultimately may look like abstract noncommittal generalities, but they can empower those who have to find the practical solutions with the right guidance for implementation.

Finally, international principles should be part of a common policy strategy to seize the new opportunities to increase the return on public investment in research and enhance the productivity and quality of research. The high-level principles should have primacy—they are the *Why* in the process. The principles then need to be implemented in a sensible access regime by the research organisations – the *How* in the process.

Contractual templates for the flexible implementation of the openness principle

To implement the general guiding principles, one way to deal with the potential imbalance in the statutory intellectual property system is to seek to amend the aspects that affect public research most negatively. However, this is not easily done, especially in view of the fact that many of these laws are quite recent and largely have ignored such considerations as they were debated and enacted.

There is, however, another and rather different approach whose practical aspects merit wide attention and support to its further development. The proposed approach consists of the voluntary use of the rights held by intellectual property owners, which allow them to construct by means of licensing contracts conditions of ‘common-use’ that emulate the key features of the public domain that are most beneficial for collaborative research in all its forms. The intention is to promote the cooperative use of scientific data, information, materials and research tools that actually are not in the public domain, and whose licensed use is therefore legally protected by an intellectual property regime. Such an undertaking may be properly described as creating ‘global information commons for science’, inasmuch as a ‘common’ constitutes a collectively held and managed bundle of resources to which access by cooperating parties is rendered open (though perhaps limited in its extent or use) under minimal transactions cost conditions.

The economic logic and practical feasibility of the ‘contractually constructed commons’ approach can be derived from non-market mechanisms constructed as systems of customary rights and restraints. Historically, it was deliberate acts of private enclosure rather than some imagined tragedy of over-grazing that often spelled the end of the agrarian commons. The legal system today makes it possible for the

owners of a tangible resource held in common to protect their collective use-rights, and manage their contractually constructed common-pool so as to sustain and augment the benefits that it yields. Consequently, because information cannot be depleted by overuse, individuals having private ownership rights in intellectual property may voluntarily use contracts to construct a common use-rights area that is all inclusive, in granting access to those wishing to use the contents. Furthermore, and because the common in this case is owned and not part of the public domain, the benefits that all users can enjoy from such an arrangement may be preserved and enhanced. This can be accomplished by reserving the legal right to exclude certain usage practices that might otherwise undermine the willingness of others to similarly pool the information that they have created.

The respective rights of the participants in the public research system can be most effectively mediated through the use of contracts at the individual researcher and institutional levels. Common-use licensing approaches that promote broad access and reuse rather than restrict it, such as those being developed by the new Science Commons under the Creative Commons mentioned ‘Emerging Open Access Models’, above, can preserve essential ownership rights while improving the social benefits and returns on the public investments in research.⁵⁰ They can help to achieve a productive balance between the domains of proprietary R&D and publicly- funded open science.

TOWARDS GLOBAL INFORMATION COMMONS FOR SCIENTIFIC DATA AND INFORMATION

The rationalisation of policies and practices across nations, institutions, and disciplines may be expected to result in much greater social and economic impact from the investment in public research overall by enabling greater access to and use of scientific data and information resources, and by facilitating interdisciplinary and international cooperation in public science and education. Because of the international scope of digital networks and research collaborations, strategic international approaches for building information commons are both necessary and desirable. In short, the adoption in recent years of

⁵⁰ See the companion article by Onsrud and Campbell in *Open Data for Global Science – Special Issue*, Paul Uhlig (ed), *CODATA Data Science Journal*, (2007).

the many innovative and promising open initiatives and common-use licensing approaches from the bottom up, coupled with the introduction of some new top-down policy proposals at the international level (at the OECD) and at the national level in several countries, make this an appropriate time to integrate these efforts.

It is for all the reasons established in this article that several international science policy organisations—CODATA, ICSU, and Science Commons—are joining efforts to launch the Global Information Commons for Science Initiative. This Initiative⁵¹ has the overall goal to accelerate the development and scaling up of open scientific data and information resources on a global basis, with particular focus on ‘common use’ licensing approaches. The specific objectives are to:

1. Improve understanding and increase awareness of the societal and economic benefits of easy access to and use of scientific data and information, especially focusing on those resulting from governmental or publicly funded research activities;
2. Promote the broad adoption of successful institutional and legal models for providing open availability on a sustainable basis and facilitating reuse of data and information;
3. Help coordinate the efforts of the many stakeholders in the world’s diverse research community who are engaged in devising and implementing effective

⁵¹ The original ideas for the Global Information Commons for Science Initiative were presented in a series of reports published at the U.S. National Academies, in a seminal article J Reichman and Paul Uhler, ‘A Contractually Reconstructed Research Commons for Scientific Data in a Highly Protectionist Intellectual Property Environment’ (Winter/Spring 2003) 66 *Law and Contemporary Problems* – Duke University School of Law 410–16 and in P David and M Spence, *Toward Institutional Infrastructures for e-Science: The Scope of the Challenges* (Report to the Joint Information Systems Committee of the Research Councils of Great Britain, Oxford Internet Institute Report No 2, 2003) <http://www.oii.ox.ac.uk/resources/publications/OIIRR_E-Science_0903.pdf>. These ideas were more fully fleshed out following an international workshop at UNESCO Headquarters in Paris on 1–2 September 2005 on the theme ‘Creating the Information Commons for Science: Toward Institutional Policies and Guidelines for Action’ (details of the Workshop rationale and proceedings, are available at: <<http://www.codataweb.org/UNESCOmtg/index.html>>). That event was organised by CODATA with the joint sponsorship of ICSU, ICSTI, INASP, UNESCO, and TWAS, and with the collaboration of the OECD.

approaches to attaining these objectives, with particular attention to the circumstances of the developing as well as developed countries.

4. Develop an online 'open knowledge environment' to promote all of the objectives of the Initiative, including providing an online collaboratory for work with different research communities to define, test, analyse, and create new knowledge about the information commons paradigm.

In our view, such an Initiative can help devise and promote new normative and legal structures for the exchange of data and information that are expected to be especially well-suited for the future conduct of collaborative research in many domains of science. By rationalising the policy and management systems in publicly funded research, the value of global digital networks and related technological advances to the progress of science can be fully realised.

NIH DATA AND RESOURCE SHARING, DATA RELEASE AND INTELLECTUAL PROPERTY POLICIES FOR GENOMICS COMMUNITY RESOURCE PROJECTS

Claire Driscoll¹

INTRODUCTION

Most observers predict significant health-related gains from genomics research. Policy and legal decisions made by government institutions, the courts and legislatures have the potential to make a significant impact on both the quantity and quality of effective and innovative healthcare-related products ultimately derived from the vibrant genomics research enterprise. In particular, the careful management of the intellectual property (IP) aspects of this promising area of research will be necessary to maximise scientific progress, provide appropriate incentives for investment, and ultimately ensure optimal public benefit.

It is the mission of the US National Institutes of Health (NIH), which is comprised of 27 individual institutes and is an agency of the US Department of Health and Human Services, to facilitate the translation of basic biomedical research discoveries into useful healthcare services and products. Within the NIH, the National Human Genome Research Institute (NHGRI) is the agency's lead entity for advancing human health through genetic research.

¹ Director of Technology Transfer for the National Human Genome Research Institute, National Institutes of Health. This chapter was first published as: 'NIH Data and Resource Sharing, Data Release and Intellectual Property Policies for Genomics Community Resource Projects' (2005) 15(1) *Expert Opinion on Therapeutic Patents* 1–8.

Through its stewardship of an array of infrastructure and research projects, including several innovative public-private consortia efforts, the NHGRI seeks to contribute to the genomic tools, data and knowledge bases. In general, I believe that scientific progress in this still young field will be best served by early, open and continuing access to: i) comprehensive, high-quality data sets containing basic biological and biochemical data; and ii) critical biological materials such as animal models and genes. Data such as the complete nucleotide sequence of many different organisms' genomes, information on genetic variation within and among populations, and results on how gene expression is regulated at the cellular and molecular level are often referred to as 'pre-competitive' information, and in my view should be made rapidly available to all, without restrictions on use. Adherence by data and resource producers and users to this simple strategy should ensure that industry and academic researchers will be able to build upon this strong foundation.

At the NIH we are expected to support basic scientific discovery whilst simultaneously facilitating the appropriate commercial research and development of the results of our formidable research programs. A sizeable number of end users for these resource projects are employed with private sector companies. For this constituency the terms governing the data use, data release, the sharing and distribution of research resources and intellectual property rights of derivative inventions are of particular importance. Policies that limit companies' ability to file patent applications or licence downstream uses could end up having an unintended inhibitory effect on the development of biomedical products. Government policies need to balance the important dual goals of continuing to rapidly place huge amounts of data in the public domain and encouraging restriction-free sharing of genomic tools, whilst also ensuring that more applied inventions, notably those closer to being an actual product, can be patented. US taxpayers, and especially patients, would like the government to appropriately foster the commercialisation of promising inventions derived from use of the data and reagents generated by these efforts. Currently, the NHGRI is actively involved in the development and vetting of policy options aimed at ensuring that genomic tools, resources and databases of genomic information are used in a manner that promotes scientific research and the practice of medicine.

Relevant policies implemented by NIH-supported public private consortia efforts such the International Human Genome Sequencing Consortium (IHGSC),² the Trans-NIH Mouse Initiative,³ the Mammalian Gene Collection (MGC)⁴ and the International Haplotype Map Project (HapMap)⁵ are specifically covered in this review.

THE NATIONAL HUMAN GENOME RESEARCH INSTITUTE: WHERE HAVE WE COME FROM AND WHERE ARE WE GOING?

Where have we come from? As leader in the international Human Genome Project (HGP), NHGRI has learned a great deal about how to coordinate and manage an international, geographically dispersed and extremely complex ‘community resource project’ (though this term was not in our lexicon at the time). A decision on how to handle the prepublication release of HGP sequence data was made early on in the project, and a policy was put in place in 1997; a year prior to the start of the major large-scale sequencing phase. The NHGRI’s data release and data deposition policy for DNA sequence information was designed ‘to make sequence data available to the research community as possible for free, unfettered use’.⁶ One of many other innovations of the HGP was the decision to include a research program on the ethical, legal and social issues (ELSI) arising from the study of the human genome. ELSI research projects, including studies on the patenting and licensing of gene patents and diagnostic tests, were funded alongside the technology development, mapping and sequencing projects.

In spring 2003 the International Human Genome Sequencing Consortium (IHGSC) celebrated the completion of the sequencing of the human genome; a milestone that, by a happy coincidence, occurred

² NHGRI, ‘International Consortium Completes Human Genome Project’ (Press Release, 14 April 2003) <<http://www.genome.gov/11006929>>; Human Genome Resources <<http://www.ncbi.nlm.nih.gov/genome/guide/human/>>.

³ Trans-NIH Mouse Initiative <<http://www.nih.gov/science/models/mouse/sharing/1.html>>.

⁴ Mammalian Gene Collection <<http://mgc.nci.nih.gov>>.

⁵ International HapMap Project <www.hapmap.org>.

⁶ NHGRI, *Policy on Release of Human Genomic Sequence Data* <<http://www.genome.gov/10000910>>.

during the 50th anniversary year of Watson and Crick's seminal discovery of the structure of DNA. In anticipation of the 'what's next' question, the NHGRI had invested in a deliberate, transparent and in-depth consultation process, lasting almost two years, with the extended genetics community about its future. With the help of patient advocates, ethicists, biotechnology company executives and healthcare practitioners, a consensus began to emerge as to which projects should be taken on by the NHGRI over the next decade.

In a paper entitled 'A Vision for the Future of Genomics Research'⁷ published in April 2003, the leaders of NHGRI outlined a series of Hilbertian⁸ grand challenges. Several of these challenges concern policy development. Both the creators and end users of the fruits of genomic research need to be actively involved in the development and vetting of policy options so that practical solutions ones that facilitate the extensive use of genomic information in laboratory and clinical settings can be implemented. Which policy issues are expected to have the greatest impact on whether or not citizens of the globe will realise benefits from genomics research in the future? Among the critical ones are data release, data and resource sharing, and patent and licensing policies. The remainder of this article focuses on the specifics of the intellectual property, data release, and research tool and data sharing policies in use for several NHGRI-funded genomics projects. Lastly, an overview of some patent and licensing-related issues that have emerged in the genomics field is provided, together with a brief summary of several possible legislative and policy fixes for these thorny IP-related problems.

GENOMICS COMMUNITY RESOURCE PROJECTS

In an important recent international meeting to discuss data release for such enterprises, the concept of a 'community resource project' was born. This is defined as 'a research project specifically devised and implemented to create a set of data, reagents or other materials whose primary utility will be as a resource for the broad scientific community'.⁹

⁷ FS Collins et al, 'A Vision for the Future of Genomics Research' (2003) 422 *Nature* 1–13.

⁸ D Hilbert, 'Mathematical Problems' (1902) 8 *Bulletin of the American Mathematical Society* 437–79.

⁹ Wellcome Trust, *Sharing Data from Large-Scale Biological Research Projects: A System of Tripartite Responsibility* (Report of a meeting organised by the Wellcome Trust, Fort Lauderdale, 14–15 January 2003).

Box 1 provides examples of several large-scale genomics enterprises that aim to produce data or create valuable scientific reagents of broad potential utility. Many of the collaborators in these projects are small- and medium-sized private sector biotechnology and biopharmaceutical companies and large global pharmaceutical giants; however, academic centers have generally played a more major role.

Box 1. Examples of Genomics Community Resource Projects

- International Human Genome Sequencing Consortium (IHGSC)/Human Genome Project (HGP):¹⁰ large-scale DNA mapping and sequencing of the human genome.
- Mouse Genome Sequencing Consortium (MGSC):¹¹ MGSC - large-scale DNA mapping and sequencing of the *Mus musculus* (mouse) genome.
- Rat Genome Sequencing Consortium (RGSC):¹² large-scale DNA mapping and sequencing of the *Rattus norvegicus* (rat) genome.
- The SNP Consortium (TSC):¹³ discovery and characterisation of two million single nucleotide polymorphisms (SNPs) in the human genome; primarily a private sector initiative with some limited NIH involvement.
- Trans-NIH Mouse Initiative:¹⁴ a group of initiatives for the creation and distribution of mouse genomic resources such as mutant mice, phenotypic and genotypic information and functional genomic data.
- Mammalian Gene Collection (MGC):¹⁵ development and distribution of a complete collection of full-length cDNAs for all known mouse, rat and human genes.

¹⁰ NHGRI, 'International Consortium Completes Human Genome Project' (Press Release, 14 April 2003) <<http://www.genome.gov/11006929>>; Human Genome Resources <<http://www.ncbi.nlm.nih.gov/genome/guide/human/>>.

¹¹ Mouse Genome Sequencing Consortium (MGSC) <<http://www.genome.gov/10001859>>.

¹² Rat Genome Sequencing Consortium (RGSC) <<http://www.genome.gov/11511308>>.

¹³ The SNP Consortium <<http://snp.cshl.org>>.

¹⁴ Trans-NIH Mouse Initiative <<http://www.nih.gov/science/models/mouse/>>.

¹⁵ Mammalian Gene Collection <<http://mge.nci.nih.gov>>.

- International Haplotype Map Project (HapMap):¹⁶ creation of a haplotype map for the comprehensive study of human DNA variation among and between a diverse set of populations.
- Encyclopaedia of DNA Elements (ENCODE):¹⁷ identification and characterisation of all the functional elements, such as regulatory sequences, encoded in the human DNA genome.

Attendees of the 1996 Bermuda¹⁸ and the 2003 Fort Lauderdale¹⁹ International Sequencing Consortium (ISC) meetings wholeheartedly agreed that the policy of rapid prepublication release of sequence data for projects, such as the HGP, should be extended beyond ‘simple’ sequence data to other types of more complex genomic data; for example, gene expression analysis/microarray data and protein structure information. ‘The products of community resource projects have, over the past several years, become increasingly important as drivers of progress in biomedical research. The scientific community will best be served if the results of community resource projects are made immediately available for free and unrestricted use by the scientific community to engage in the full range of opportunities for creative science.’²⁰ The conclusion reached was that the architects of these other large-scale projects should seriously consider adopting the same data release policy. Of course, the exact details of how, in what format and under what type of schedule these more complex data will be downloaded into public databases still needs to be defined by the domain experts. For example, there are problems related to the optimal level and degree of validation needed for particular kinds of experimental data so that the data are useful and useable to other consortia scientists and non-consortia scientists.

¹⁶ International HapMap Project <www.hapmap.org>.

¹⁷ Encyclopaedia of DNA Elements (ENCODE) <<http://www.nhgri.nih.gov/10005107>>.

¹⁸ Dr Bentley, ‘Genomic Sequence Information Should be Released Immediately and Freely in the Public Domain’ (1996) 274 *Science* 533–4; Summary of Principles agreed at the International Strategy Meeting on Human Genome Sequencing (Bermuda, 25–28 February 1996).

¹⁹ Wellcome Trust <<http://www.wellcome.ac.uk/en/1/awtpubrepdat.html>>.

²⁰ Wellcome Trust <<http://www.wellcome.ac.uk/en/1/awtpubrepdat.html>>.

The conclusions reached above by the ISC participants are consistent with the now seven-year-old NIH research tools policy officially called ‘Sharing Biomedical Research Resources: Principles and Guidelines for Recipients of NIH Research Grants and Contracts’²¹ and the policies are also in line with the newer NIH policy on the sharing of data for extramural grantees and contractors.²² It is important to note that databases, along with research reagents, such as genes, vectors and antibodies, are considered to be research tools. Under the policy, which went into effect on October 1, 2003, grantee institutions must submit proposed data-sharing plans in all grant applications that request US \$500 000 or more per year in funds.

SPECIAL CHALLENGES IN ACHIEVING OPEN DATA ACCESS: THE SNP CONSORTIUM AND THE HAPMAP PROJECTS

The SNP Consortium (TSC) is a non-profit foundation that focused on discovering single point mutations in the human genome called single nucleotide polymorphisms (SNPs). It was formed in 1999 with a mix of public and private funds, and the final data release occurred last September. Its mission was to ‘develop up to 300 000 SNPs distributed evenly throughout the human genome and to make the information related to these SNPs available to the public without intellectual property restrictions’.²³ The main idea behind the unique structuring of the consortium was that it made sense to have all interested parties, which included large pharmaceutical companies such as GlaxoSmithKline and charities such as the Wellcome Trust, share the risk and expense of developing a high-quality publicly available human SNP data set. Many companies participated even though they were not given any special use of, or early access to, the data. Nonetheless, clearly the private sector felt that the effort was worth the investment.

The TSC members agreed to adopt a policy of waiving the right to receive patent protection on the raw SNP data and agreed to publish the

²¹ *Report of the NIH Working Group on Research Tools*, presented to the Advisory Committee to the Director (4 June 1998) <<http://www.nih.gov/news/researchtools/>>.

²² NIH, *Data Sharing Information* <http://grants2.nih.gov/grants/policy/data_sharing/>.

²³ The SNP Consortium <<http://snp.cshl.org>>.

mapped SNPs as quickly as was feasible. A legal instrument called a Statutory Invention Registration (SIR) was used in a defensive tactic to guard against those who might try to file patent applications on identical SNP data in advance of its public deposition. SIRs were filed on the data as it was generated and then those SIRs were abandoned once the validated SNP information was downloaded to the TSC website every three months.²⁴

The goal of the HapMap consortium-based effort is to determine the common patterns of sequence variation in the human genome.²⁵ Because *de novo* mutations occur relatively rarely in human populations any new allele that does arise will ‘travel’ with other nearby alleles within a continuous block of genomic DNA on the particular chromosome. These common sets of alleles are called haplotypes. At present the genotyping of hundreds of samples and the construction of detailed haplotype maps is being carried out on set of DNA specimens from populations with ancestry from various regions in Asia, Africa and Europe.²⁶

‘The [HapMap] project is committed to rapid and complete data release, and to ensuring that project data remain freely available in the public domain at no cost to users.’²⁷ A legal review of potential intellectual property problems with the project raised one major concern: it would be theoretically possible for an unscrupulous company or entity to add on a trivial amount of information to the published HapMap data and then attempt to secure ‘parasitic’ patent claims such that all others would be prohibited from using the original public data. This scenario along with other related concerns and a desire to be in harmony with earlier community resource projects, led the consortium members to agree upon a new data release and data use policy. Under the policy, users of HapMap data can file patent applications on associations they uncover or verify between particular SNPs and disease and/or disease susceptibility. The only caveat is that the owner or assignee of the

²⁴ The SNP Consortium, *Frequently Asked Questions* <<http://snp.cshl.org/about/faq.shtml>>.

²⁵ International HapMap Project <www.hapmap.org>.

²⁶ The International HapMap Consortium, ‘International HapMap Consortium Paper’ (2003) 426 *Nature* 789–96.

²⁷ The International HapMap Consortium, ‘International HapMap Consortium Paper’ (2003) 426 *Nature* 789–96.

patents cannot try to limit or prevent other users from enjoying full and unrestricted access to the HapMap data. A click-on licence agreement describing the details of the policy is posted on the HapMap project website.²⁸ Users must agree to the conditions of the online policy before they are permitted to access or download HapMap data.

The HapMap participants, similarly to the members of TSC, agreed that SNPs, genotypes and haplotype data of unknown specific utility are not inventions and therefore are not appropriate subject matter for patenting.²⁹ Conversely, an SNP or haplotype that is strongly associated with a disease or medically important phenotype, such as susceptibility to diabetes or a poor response to a particular chemotherapy drug, would clearly have a specific utility and be patentable. However, neither TSC nor the HapMap projects include phenotype or disease association studies.

THE SHARING OF VALUABLE RESEARCH REAGENTS: MOUSE RESOURCES AND THE MAMMALIAN GENE COLLECTION

The NIH encourages and actively supports the sharing of resources such as inbred mouse strains, genetically modified mice as well as the DNA vectors and murine embryonic stem cells that must be used to generate useful model knockout lines. To facilitate the timely and efficient development and distribution of a wide array of mouse resources the NIH decided to place a number of related projects under the administrative management of a new umbrella program called the Trans-NIH Mouse Initiative.³⁰ Beginning in 1998, workshops were convened to bring together diverse members of mouse and larger scientific community, as well as program staff from many NIH institutes, in order to research agreement on the key priorities for the initiative. Coordination and oversight of the many program such mutagenesis and phenotyping studies, genomic sequencing and the creation of mouse repositories such as the Mutant Mouse Regional Resource Centers is

²⁸ International HapMap Project <www.hapmap.org>.

²⁹ International HapMap Project <www.hapmap.org>.

³⁰ Trans-NIH Mouse Initiative <<http://www.nih.gov/science/models/mouse/>>.

carried out by a Trans-NIH Mouse Genomics and Genetics Resources Coordinating Group.

In 2003, the coordinating group also wrote a policy document entitled 'NIH Statement on Sharing and Distributing Mouse Resources'.³¹ As a result of this policy, not-for-profit entities may obtain materials for use in non-commercial research using a standard Material Transfer Agreement (MTA). Importantly, the MTA cannot contain reach-through licensing terms, nor can it contain provisions that are more restrictive than those included in the widely used Simple Letter Agreement (SLA)/Uniform Biological Materials Transfer Agreements (UBMTA).³² In most cases the decision to seek patent protection on a particular genetically modified mouse is entirely up to the grantee or contractor, as is their right under the US 1980 *Bayh Dole Act*.

The terms of some NIH funding arrangements for projects under the Trans-NIH Mouse Initiative do, however, contain an exception, known as a Determination of Exceptional Circumstances (DEC). US government grantees and contractors do not obtain title to inventions when the government makes a 'determination of exceptional circumstances'.³³ A DEC removes the standard right of ownership usually retained by the funded institution for inventions made by their employees when the research was funded, partially or in whole, with government monies. This administrative and legal tool is only used for programs in which the main goals are to create data and/or resources that are to be made widely available with minimal restrictions on their use. NIH's intention to utilise a DEC for a new program is always announced in advance of the award of a grant or contract. Regardless of whether or not a specific mouse mutant or mouse resource is patented those interested in using the animals or materials for commercial purposes must negotiate a licence with the owner.

Another important genomics research tool initiative called the Mammalian Gene Collection (MGC) program, which began in 1999, represents yet another unique undertaking by NIH. Its remit is to create

³¹ NIH, *Statement of Sharing and Distributing Mouse Resources* <<http://www.nih.gov/science/models/mouse/sharing/>>.

³² Uniform Biological Materials Transfer Agreement (UBMTA) <<http://ott.od.nih.gov/NewPages/UBMTA.pdf>>.

³³ As set forth at: 35 USC § 202(a) (1994).

a public collection of affordable, sequence-verified full-length complementary DNAs (cDNAs) for every known mouse and human gene, and a subset of rat genes. Feedback solicited from the NIH intramural and extramural communities at the time revealed that the lack of reasonably priced, freely available and sequence-verified ‘correct’ cDNA clones was indeed slowing down many research projects. After a careful study and review of the various options, the NIH decided to fund and handle the coordination for the MGC initiative. This was fortunate as perhaps no other entity had the experience, financial resources or technical wherewithal to competently take on and successfully complete such a project.

Like some programs managed under the Trans-NIH Mouse Initiative, the MGC request for funding proposals contained a DEC. The MGC also works closely with related projects to create collections of *Xenopus* and zebrafish full-length cDNAs. Not only does the MGC make available high- quality and modestly priced cDNAs but buyers are free to use the clones for research purposes and there are no onerous IP or reach-through licensing terms whatsoever.

Today the MGC physical clones are currently available to both non-commercial and commercial scientists via the IMAGE consortium distributors for a modest fee.³⁴ In addition, as a result of NIH partner Invitrogen’s open architecture licensing policy for its Gateway® and Superscript technologies, as is also outlined in an open-access agreement with NIH, academic and government purchasers of the system are permitted to share MGC clones made using the company’s technologies and/or genes cloned into Gateway® vectors with others for research purposes. The HGP, the HapMap, the Trans-NIH Mouse Initiative and the MGC are key components of the burgeoning publicly available collections of scientific resources (mainly data and research reagents) that are supported by NHGRI and other NIH Institutes. The creation of additional valuable research reagents and data sets, available to all, for the global genomics-based research toolbox is a goal that cross-cuts all of NHGRI’s community resource projects.

³⁴ Mammalian Gene Collection, *Where to Buy* <<http://mgc.nci.nih.gov/Info/Buy>>.

NIH PATENT POLICY AND GENOMICS-SPECIFIC INTELLECTUAL PROPERTY ISSUES

As a result of the 1986 *Federal Technology Transfer Act (FTTA)* American government research laboratories such as NIH were given a statutory mandate to encourage and facilitate the efficient and expeditious transfer of promising new technologies invented in NIH intramural labs to companies for further development and commercialisation. Over the next few years government agencies created technology licensing offices and began to develop and implement patent and licensing policies. The current version of the NIH Patent Policy was adopted in 1995 and includes the following recommendations:³⁵

- Seek patent protection on biomedical technologies when having IP rights will facilitate the availability of the technology for research or commercial use.
- Seek patent protection for inventions when IP rights are necessary to attract commercial partners and further R&D is required.
- Do not seek patent protection for technologies if no further R&D is required and the invention has no obvious preventative, diagnostic/prognostic and/or therapeutic use (for example, has no commercial or public health value).
- Do not unduly delay or avoid the public disclosure of research results because of anticipated future patent filings (reasonable delays of one to several months are permitted).

The above policy does not distinguish between different types of inventions; it is purposefully general in nature so that it can be applied to all NIH-developed inventions.

Earlier in 2004, the NIH opted to publish a guidance document specifically aimed at outlining what are considered to be our own internal best practices for the handling of genomic inventions. This new draft publication was entitled 'Best Practices for the Licensing of

³⁵ *United States Public Health Service Technology Transfer Manual*, Chapter 200: PHS Patent Policy.

Genomics Inventions',³⁶ and it summarised the guiding principles used by licensing specialists at the NIH when they negotiate deals. One of the most provocative, and in my view the most sensible, suggestions had to do with the licensing of diagnostic applications of genomic inventions: entities are strongly encouraged to non-exclusively licence diagnostic inventions or, at a minimum, to grant only narrow exclusive licences for these kinds of inventions (limiting the licence to particular fields of use such as one type of technology platform, one or a few disease indication(s) and/or certain geographic regions). These suggestions are similar to those put forward by the Nuffield Council on Bioethics in their 2002 discussion paper.³⁷

It is important to encourage the commercialisation of as many types of clinical diagnostic tools and tests as possible, as competition in the marketplace should translate to an increase in the number of innovative and affordable products available to patients and their physicians. By publishing practical details on how we at NIH handle the licensing of various types of genomic inventions, we hope to give academics ideas for how they might handle the licensing of similar inventions within their own universities and hospitals.

To get a patent in the US one must show that the invention is: useful, non-obvious or inventive, and novel. The Nuffield Council on Bioethics report³⁸ on the ethics of patenting DNA discusses several important issues with regard to the legal concepts of utility and inventiveness. In general, the European Patent Office (EPO) holds patent applications to a higher standard of inventiveness than does the US Patent and Trademark Office (USPTO). The EPO states 'the isolation of DNA sequences that have a structure closely related to existing sequences in which the function is known is not inventive'.³⁹ Using computational tools to identify homologues and assign hypothetical functions, so-called *in silico* analysis, to a DNA sequence

³⁶ *Best Practices for the Licensing of Genomic Inventions*, 69 Federal Regulations 67747–8 (2004).

³⁷ Nuffield Council on Bioethics, *The Ethics of Patenting DNA: A Discussion Paper* (July 2002) <www.nuffieldbioethics.org>.

³⁸ Nuffield Council on Bioethics, *The Ethics of Patenting DNA: A Discussion Paper* (July 2002) <www.nuffieldbioethics.org>.

³⁹ Nuffield Council on Bioethics, *The Ethics of Patenting DNA: A Discussion Paper* (July 2002) <www.nuffieldbioethics.org>.

would not be considered inventive in Europe; whereas according to the USPTO rules the use of non-laboratory bench computer-based methods would not necessarily exclude the granting of a patent for a gene on the basis of non-obviousness.⁴⁰ Fortunately, the USPTO issued revised utility guidelines in early 2001. Inventions now must show a 'specific and substantial and credible utility'.⁴¹ This is certainly an improvement but only one specific, substantial and credible utility per DNA sequence is needed and so a composition-of-matter patent on a gene sequence can easily come to encompass new uses discovered long after the initial filing, even if those uses were not known at the time or not disclosed in the original patent.⁴² In my opinion this is unfair as it penalises those who do the hard work of figuring out the biological function(s) of genes and proteins. The USPTO utility guidelines should be modified to eliminate this 'loophole'.

Box 2 contains information on three published reports and one ongoing study, all of which focus on gene patent and/or genomic patent and licensing issues. Patents are a lynchpin for successful commercial entities in the genomics field; they are a driving force behind innovation and a guarantee that the discoveries will be fully disclosed and speedily delivered for scientific and societal benefit. Nonetheless, patents with overly broad claims, as well as those with unsubstantiated data or a complete lack of data for the specific claimed uses, should not be allowed to issue. Also, there is the issue of the exponentially growing body of patents in the genomics area; this has led to a so-called patent 'thicket'⁴³ problem. To ensure their freedom to operate in a complex marketplace many companies often must licence a large range of overlapping and related patents; the high licensing and transactional costs in terms of both time and money could be causing some companies to abandon efforts to try to commercialise needed diagnostic and therapeutic products.

⁴⁰ Nuffield Council on Bioethics, *The Ethics of Patenting DNA: A Discussion Paper* (July 2002) <www.nuffieldbioethics.org>.

⁴¹ *USPTO Utility Examination Guidelines*, 66 Federal Regulations 1092 (2001).

⁴² E Marshall, 'Patent on HIV Receptor Provokes an Outcry' (2000) 287 *Science* 1375–7.

⁴³ MA Heller and RS Eisenberg, 'Can Patents Deter Innovation? The Anticommons in Biomedical Research' (1998) 280 *Science* 698–701.

Box 2. Recent Gene Patent and Intellectual Property Policy Reports

- Nuffield Council on Bioethics, *The Ethics of Patenting DNA: a Discussion Paper* (July 2002) <www.nuffieldbioethics.org>.
- The Royal Society, *Keeping Science Open: the Effects of Intellectual Property on the Conduct of Science* (April 2003) <www.royalsoc.ac.uk>.
- The Australian Law Reform Commission, *Gene Patenting and Human Health Discussion Paper* (March 2004) <www.austlii.edu.au/au/other/alrc/publications/dp/68/>.
- US National Academy of Sciences, *Intellectual Property Rights in Genomics and Protein-Related Research* (Report expected in March 2005) <www.nationalacademies.org> (under Current Projects tab).

The sequencing and public disclosure of the human genome and the gearing up of several other genomics community resource projects has effected a shift in the commercial landscape. Over the last few years, with sequence information, SNP information and soon haplotype data publicly available at no cost, a number of companies have exited the genomic information database subscription business. Not surprisingly, many of these same ‘content’ companies are in the process of transforming themselves into biopharmaceutical companies. Consequently, they are now focusing exclusively on activities such as functional genomics, target validation and the screening of small molecular libraries in their quest to identify promising lead molecules.

Finally we must ponder the ramifications of a recent US court ruling. Historically, biotechnology and pharmaceutical companies have almost never sued academic investigators for patent infringement, as long as no active selling was occurring, even if the not-for-profit scientists were using the company’s patented invention. The harsh reality of the lack of a formal research exemption in US law was brought to everyone’s attention with the *Madey v Duke*⁴⁴ court decision. As succinctly stated by

⁴⁴ *Madey v Duke University* 307 F 3d 1351 (2002).

University of Michigan law professor Rebecca Eisenberg, '[t]he recent rejection by the Court of Appeals for the Federal Circuit of an "experimental use defence" to a patent infringement lawsuit against Duke University. . . is an alarming wake-up call to the academic community'.⁴⁵ That court decision now makes it possible, and perhaps even likely, that cash-strapped or aggressive companies with no immediate revenue streams will assert their patent rights against university researchers.⁴⁶ Rather than taking a wait-and-see approach, an attempt should be made by lawmakers and academics to create a true research exemption in US patent law.

EXPERT OPINION

The free sharing and wide dissemination of pre-competitive genomic data and research resources has numerous significant benefits to research institutions, companies, scientists and the general public. Therefore, we do not support the idea of the US enacting *sui generis* database protection legislation similar to the Database Directive passed in the EU.⁴⁷ Share and share alike 'open source code' - like data release and use policies such as the ones mentioned in this article inevitably encourage researchers to undertake different scientific approaches as they attempt to unravel the intricacies of complex biological systems. The following are just a few of the benefits that result from the creation and maintenance of centralised databases and repositories: improved ease of access; avoidance of duplication of effort; and more efficient use of limited research funds. Projects such as the IHGSC, International Mouse Genome Sequencing Consortium (IMGSC), International Rat Genome Sequencing Consortium (IRGSC), TSC and HapMap have unequivocally demonstrated the usefulness of having open access comprehensive databases that contain vast quantities of genomic information.

Initiatives such as the MGC and Trans-NIH Mouse Initiative have drawn attention to the immense value of research tools for aiding and

⁴⁵ RS Eisenberg, 'Patent Swords and Shields' (2003) 299 *Science* 1018–19.

⁴⁶ C Dennis, 'Geneticists Question Fees for Use of Patented "Junk" DNA' (2003) 423 *Nature* 105.

⁴⁷ EU Database Directive 96/9/EC (1996)

<<http://europa.eu.int/ISPO/infosoc/legreg/docs/969ec.html>>.

accelerating scientific progress in all sectors: academic, government and commercial. All of these projects bring together the collective financial and intellectual resources of many diverse partners, and result in efficiencies of operation, scale and speed not normally associated with government- or academic-backed endeavours. Participants in Genomics Community Resource Projects are cognisant that the greatest opportunity for value creation using the data and research tools developed by consortia efforts will occur downstream from the discovery of basic genetic and biological information.

CONCLUSION

In summary: i) I hope that universities and others will follow the suggestions outlined in NIH's new guidance document on the best practice for licensing genomic inventions; ii) I support the concept of enacting legislation to establish a formal research exemption for the non-commercial research use of patented inventions and technologies by not-for-profit entities; iii) I support the idea of raising the utility 'bar' even higher than was done in the 2001 revisions of the USPTO's utility guidelines for biotechnology inventions; and iv) I support the concept of compulsory licensing and/or compulsory sublicensing by patent holders of genomic/genetic diagnostic technologies and inventions.⁴⁸

Today it is much easier for scientists everywhere to rapidly build upon genomic research carried out by their colleagues. We at NIH look forward to the establishment of additional public private partnerships that are dedicated to placing pre-competitive data in the public domain. As a leader in the fields of genetic research and genomic science, NHGRI will continue to help encourage policy debates and support research,⁴⁹ which informs policy development, in order to facilitate the widespread use of genomic information in both research and clinical settings, and in order to facilitate the development of new genomics-based products.

⁴⁸ Proposed Genomic Research and Diagnostic Accessibility Act of 2002 (HR3967); T Abate, 'Do Patients Wrap Research in Red Tape' *San Francisco Chronicle*, 25 March 2002.

⁴⁹ Example: request for application RFA-HG-04-004: *Intellectual Property Rights in Genetics and Genomics* <<http://grants1.nih.gov/grants/guide/rfa-files/RFA-HG-04-004.html>>.

CYBERINFRASTRUCTURE FOR KNOWLEDGE SHARING

John Wilbanks¹

Infrastructure never gets adequately funded because it cuts across disciplinary boundaries, it doesn't benefit particular groups. Infrastructure is a prerequisite to great leaps forward and is thus never captured within disciplinary funding, or normal governmental operations. We need to revise radically our conception of cyberinfrastructure. It isn't just a set of tubes through which bytes flow, it is a set of structures that network different areas of knowledge ... and that is software and social engineering, not fiber optic cable. The superhighways of the biological information age should not be understood as simply physical data roads, long ropes of fiber and glass. They need to be structures of knowledge. The Eisenhower Freeways of Biological Knowledge are yet to be built. But that doesn't mean the task isn't worth starting.

– James Boyle, William Neal Reynolds Professor of Law,
Duke University Law School

KNOWLEDGE SHARING AND SCHOLARLY PROGRESS

Knowledge sharing is at the root of scholarship and science. A hypothesis is formulated, research performed, experimental materials designed or acquired, tests run, data obtained and analysed, and finally a publication. The scholar writes a document outlining the work for dissemination in a scholarly journal.

¹ Executive Director of Science Commons. This chapter was first published as an article in (2007) 3 (3) *CTWatch Quarterly* <<http://www.ctwatch.org/quarterly/articles/2007/08/cyberinfrastructure-for-knowledge-sharing/>>.

If it passes the litmus test of peer review, the research enters the canon of the discipline. Over time, it may become a classic with hundreds of citations. Or, more likely, it will join the vast majority of research, with less than two citations over its lifetime, its asserted contributions to the canon increasingly difficult to find – because, in our current world, citations are the best measure of relevance-based search available.

But no matter the fate of an individual publication, the system of publishing is a system of sharing knowledge. We publish as scholars and scientists to share our discoveries with the world (and, of course, to be credited with those discoveries through additional research funding, tenure, and more). And this system has served science extraordinarily well over the more than three hundred years since scholarly journals were birthed in France and England.

THE INFORMATION TECHNOLOGY REVOLUTION: MISSED CONNECTIONS AND LOST OPPORTUNITIES

Into this old and venerable system has come the earthquake of modern information and communication technologies. The Internet and the Web have made publication cheap and sharing easy – from a technical perspective. The cost of moving, copying, forwarding, and storing the bits in a single scientific publication approach zero.

These technologies have created both enormous efficiency gains in traditional industries (think about how Wal-Mart uses the network to optimise its supply chains) and radical reformulation of industry (Amazon.com in books, or iTunes in music). Yet the promise of enormous increases in efficiency and radical reformulations have to date failed to make similar shattering changes to the rate of meaningful discovery in many scientific disciplines.

For the purposes of this article, I focus on the life sciences in particular. The problems I articulate affect all the scientific disciplines to one extent or another – but the life sciences represent an ideal discussion case. The life sciences are endlessly complex and the problems of global health and pharmaceutical productivity such an enormous burden that the pain of a missed connection is personal. Climate change represents a problem of similar complexity and import to the world, and this article should be contemplated as bearing on research there as well, but my topic is in the

application of cyberinfrastructure to the life sciences, and there I'll try to remain.

Despite new technology after new technology, the cost of discovering a drug keeps increasing, and the return on investment in life sciences (as measured by new drugs hitting the market for new diseases) keeps dropping. While the Web and email pervade pharmaceutical companies, the elusive goal remains 'knowledge management': finding some way to bring sanity to the sprawling mass of figures, emails, data sets, databases, slide shows, spreadsheets, and sequences that underpin advanced life sciences research. Bioinformatics, combinatorial drug discovery, systems biology, and an innumerable number of words ending with '-omics' have yet to relieve the skyrocketing costs and increase the percentage of success in clinical trials for new drug compounds.

The reasons for this are many. First and foremost, drug discovery is hard – really, really hard. And much of the low-hanging fruit has been picked. There are other reasons having to do with regulatory requirements, scientific competition, distortions in funding, and more. But there is one reason that stands out as both a significant drag on discovery and as a *treatable* problem, one that actually can be solved in the short term: we aren't sharing knowledge as efficiently as we could be.

FORGET 'WEB 2.0' – WHAT ABOUT 'WEB 1.0' FOR SCIENCE?

Much of the functionality we take for granted on the Web comes from making the choice to make sharing information easier, not harder. A good example is the way that Google interacts with the scientific literature.

With few exceptions, we rank the importance and relevance of scientific articles the way we always have, with citations and 'impact factors'. Citations are longstanding and important. Impact factors – the number of citations to the articles in a journal – are the dominant metric for journal quality. And for a long time, citations were clearly the best, and perhaps the only, statistical measure of quality of a journal. In a print world, a world without hyperlinks and search engines and blogs and collaborative filtering, citations are a beacon of relevance.

But we live in a different world now. We have the ability to make connection after connection between documents, to traverse easily from

one page to another page. Hyperlinks are cheap and they're everywhere. It was a conscious design decision made by Tim Berners-Lee to allow this functionality. Other competing systems thought it insane that the WWW would let just anyone link to just anything else – those links might be broken, leading to the dreaded '404 not found' – and that would obviously kill the WWW! It hasn't worked out that way. The choice to allow users the right to make hyperlinks, to make hyperlinking easy and fast, not only did not kill the Web, it is a big part of what makes Google searching so powerful.

Google ranks pages by downloading enormous chunks of the Web and running software that analyses the linkages between Web pages. The system quite literally depends on there being lots and lots of links, many of them perhaps useless on their own, but which in aggregate provide hints of relevance. Thus, the number one Google search on the words 'Science Commons' is the Web page analysed with the words 'Science Commons' that has the most links pointing to it. There's more complexity, obviously, but that's a big part of the idea.

If those Web pages were private, the page ranking system wouldn't work. The Web pages themselves are part of the infrastructure on which Google operates, on which millions of start-up dreams are founded. In a world where every page was locked, where every Web designer had to ask permission to make an inbound link ... we wouldn't have the sprawling value creation we associate with the Internet. It would look a lot more like Prodigy looked a long time ago: a closed network that can't compete in the end with the open networks.

Put another way, we have far more efficiency brought to bear on accelerating our capability to order consumer products than we do on accelerating our capability to perform scientific research. Biological reagents and assays are re-invented and reverse-engineered by readers of 'papers' – years of laboratory work, data, living DNA and more compressed down to the digital equivalent of a sheet of dead tree.

We need the Web to work as well for science as it does for other areas. The capabilities now exist to integrate information, data, physical tools, order fulfilment, overnight shipping, online billing, one-to-one orders, and more. If we are to solve the persistent health problems of the world, of infectious disease in the developing world and rare disease in

the developed world, the 'Web 1.0' efficiency is an obvious benefit to bring to the life sciences.

But these advances we take for granted in daily life, like Google's relevance based search of the entire Web, eBay's many-to-many listing and fulfilment, Amazon's one-click ordering, won't come to science accidentally. There's a significant collective action problem blocking the adoption of these systems and preventing the network effects from taking over in discovery.

But it's not just the Google issue, which simply forces us to forego existing technology and focus on citations as we have always done. Citations carry more constrictions as a search metaphor. You are likely to enter the citation search ranked world when you know what to search. But you might not know what you're looking for. You might not know how to say it in the nomenclature of a related, but distinct, discipline.

It goes on. Citation linkages between papers are subject to enormous social pressures. One cites the papers of one's bosses, of course. Review articles can skew impact factors. And of course, a tried-and-true way to get a heavily cited article remains to be horrifically, memorably wrong.

And over the long term, the lack of more complex and realistic interconnections between articles – a web, a set of highways, an infrastructure connecting the knowledge – is that we can't begin to integrate the articles with the databases. That's because the actors in the articles (the genes, proteins, cells and diseases) are described in hundreds of databases.

And if we could link the articles not just to each other by a richer method than citations, but to the databases, we can inch closer to the goal of a Rosetta Stone of knowledge, the small element upon which we can begin to have truly integrated, public knowledge spaces. That would in turn allow us to begin automatically indexing the data that robots are producing in labs every day, to meaningfully extract actionable information from the terabytes of genomic data we are capable of producing.

You get those virtues only where you are dealing with the knowledge claims themselves, not the sub-component of them the people in the field thought it worthy to expose. Only a better infrastructure gets you there, just as the modern highway system in the US allowed for better

efficiency than the evolved hodge-podge of state highways. Citation linkages are very useful (and a later version needs to cross reference them with these highways we propose – we didn’t throw away the state highways, after all!). This is simply a different set of tasks, and one that can be accomplished if enough smart people have enough rights and time to work on the knowledge.

But sadly, no one – no one! – has the right to download and index with scholarly literature without burning years of time and money in negotiations. Google has spent years asking for the right to index a lot of the scholarly canon for its Scholar project, but that’s not some open land trust for any researcher to work on. It’s just for Google. And the fact that Google alone has the right to index articles for such a service means that the *next* Google, the next set of genius entrepreneurs with a taste for search coding away in the halls of the local university, can’t apply their skills to the sciences.

Though we have the capability to drastically increase the sharing at a much lower cost through digital distribution, search, and more, the reaction has been instead to segregate knowledge behind walls of cost, technology, and competitive secrecy. The net result is that we’re doing things the way we always did, but only somewhat faster. If we want to bring both efficiency gains and radical transformation to the life sciences, getting more knowledge online, with the rights to transform, twist, tag, reformat, translate, and more, is going to be part of the solution. We have to start allowing the best minds of the world to apply the newest technologies to the scientific problems facing us.

There isn’t a single, open ‘Web’ of content to search – it’s owned by a group of publishers who prevent indexing and search outside their own engines, and who use copyright and contracts to keep it locked up. There isn’t any easy way to find the tools of biological science – it’s a complicated social system of call-and-response, of email and phone calls, of ‘are you in the club of scientists worth partnering with?’ questions and answers. And there isn’t a standard way to get your orders fulfilled, but instead a system in itself of materials transfer and ordering, university technology transfer, commercial incentives, deliberate withholding, and more. We don’t have the Web working yet for science.

INFRASTRUCTURE FOR KNOWLEDGE SHARING: SCIENCE COMMONS

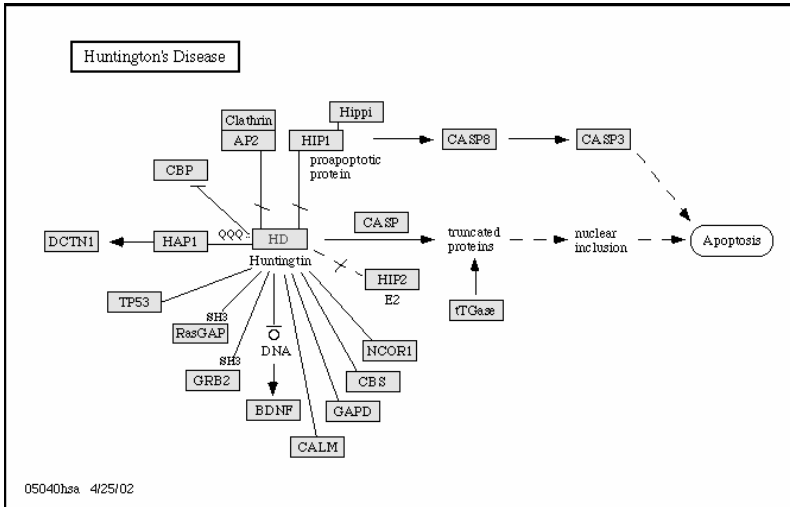
I work on a project called Science Commons – part of the Creative Commons (CC) non-profit organisation (CC is the creator of a set of legal tools for sharing copyrighted works on the Web using a modular set of machine-readable contracts. CC licenses cover more the 150 million copyrighted objects on the Web, including such high-impact offerings as BioMed Central, Public Library of Science, Nature Precedings, Hindawi Publishing, and the UniProt database of proteins. Science Commons is building a toolkit of policy, contracts, and technology that increases the chance of meaningful discovery through a shared, open approach to scientific research. We're building part of the infrastructure for knowledge sharing, and we're also deploying some test cases to demonstrate the power of investing in this kind of infrastructure.

Science Commons isn't alone. Sharing approaches that address a single piece of the research cycle are making real, but painfully slow, progress. Open Access journals are far from the standard. Biological research materials are still hard to find and harder to access. And while most data remains behind the firewall at laboratories, even those data sets that do make it online are frequently poorly annotated and hard to use. The existing approaches are not creating the radical acceleration of scientific advancement that is made possible by the technical infrastructure to generate and share information.

Science Commons represents an integrated approach – one with potential to create this radical acceleration. We are targeting three key blocking points in the scientific research cycle – access to the literature, experimental materials, and data sharing – in a unified approach. We are testing the hypothesis that the solutions to one problem represent inputs to the next problem, and that a holistic approach to the problems discussed here potentially benefits from network effects and can create disruptive change in the throughput of scientific research. I will outline how these approaches represent tentative steps towards open knowledge infrastructure in the field of neuroscience.

KNOWLEDGE OVERLOAD FOR HUNTINGTON'S DISEASE

Figure 1. Biological pathway for Huntington's Disease.



Source: Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S., Tokimatsu, T., and Yamanishi, Y., "KEGG for linking genomes to life and the environment." *Nucleic Acids. Res.* 36, D480–D484 (2008).

Above is the biological pathway for Huntington's Disease. This pathway is like a circuit – it governs the movement of information between genes and proteins, processes and locations in the cell. This one is a relatively simple pathway, as far as such things go. More complex pathways can have hundreds of elements in the network, each 'directional' - not just linked like Web pages, but typed and directed links, where the kind of relationship and the causal order are vital both in vitro and in silico.

In this pathway, the problem is the HD gene in the middle of the circuit - if that gene is broken, it leads to a cascade that causes a rare, fatal disease where the brain degenerates rapidly. Although the genetic element has been understood for a long time, there is no cure. Not enough people get the disease for it to be financially worth finding a cure, given how expensive it is to find drugs and get them to market.

That's cold comfort to the tens of thousands of people who succumb each year and to their families who know they have a 50% chance of passing on the gene and disease to their children. But that's the reality.

Years of research have led to an enormous amount of knowledge about Huntington's. For example, a search in the U.S. government's free Entrez web resource on 'Huntington's' yields more than 6000 papers, 450+ gene sequences, 200+ protein sequences, and 55 000 expression and molecular abundance profiles. That's a lot of knowledge. The papers alone would take 17 years to read, at the rate of one paper per day (and that's assuming no new papers are published in the intervening years). Yet Huntington's actually provides a relatively small result. One of the actors in the pathway is called 'TP53'. That brings up another 2500 papers, but also brings up (in an indirect link to a page about sequences for this entity) that it has a synonym: 'p53'. Entrez brings back 42 000 articles from that search string – 115 years to read!

It goes on and on. And having all of this knowledge is wonderful. But there are more than a few problems here. The first is something you might call 'cognitive overload'. Our brains simply aren't strong enough to take in 500 000 papers, read them all, build a mental model of the information, and then use that information to make decisions - decisions like, what happens if I knock out that CASP box in the pathway, with 27 000 papers?

The other problems stem from the complexity of the body. In what other circuits is each entity in the pathway involved? What about those tricky causal relationships above and below it in the circuit? What are the implications of intervention in this circuit on the other circuits?

Some of these entities, the boxes in the diagram, are metaphorically similar to the airport in Knoxville, TN. Knocking out that airport doesn't foul up a lot of air traffic. But some of these - P53 for example - are more like Chicago. Interfering with that piece of the network reverberates across a lot of unrelated pieces of the network. That's what we call side effects, and it's one of the reasons drugs are so expensive - we know that we can impact this circuit, but we don't realise how badly it affects everything else until we run the drug in the only model available that covers all possible impacts: the human body.

And this is just the papers. There are thousands of databases with valuable information in them. Each of them has different access

privilege conditions, different formats, different languages, and different goals; wasn't designed to work with anything else; and is maintained at different levels of quality. But they have vital - or potentially vital, to the right person asking the right question - information. And if we could connect the knowledge around these knowledge sources into a single network we just might be able to leverage the power of other technologies built for other networks. (Like Google – but maybe more like the next Google, something as dramatically better and different and radical as Google was when we first saw it in the late 1990s.)

There are two problems to be addressed here. One is the materials that underpin this knowledge, these databases and articles. Those materials are 'dark' to the Web, invisible, and not subject to the efficiency gains we take for granted in the consumer world. The second is the massive knowledge overload that the average scientist faces. I'll outline two proofs of concept to demonstrate the value of investment in infrastructure for knowledge sharing that can address these problems.

PROOF OF CONCEPT: E-COMMERCE FOR BIOLOGICAL MATERIALS

The Biological Materials Transfer Agreement Project (MTA) develops and deploys standard, modular contracts to lower the costs of transferring physical biological materials such as DNA, cell lines, model animals, antibodies and more. Materials represent tacit knowledge – generating a DNA plasmid or an antibody can take months or years, and replicating the work is rarely feasible. Gaining access to those materials is subject to secrecy, competition, lack of resources to manufacture materials, lack of time, legal transaction costs and delays, and more.

There is significant evidence that the transfer of biological materials is subject to significant slowdowns. Campbell² and Cohen³ have each demonstrated that materials are frequently denied. Legal barriers are part of the problem – more so than patents – but the greater problem is frequently the competition, secrecy, and incentive systems involved.

² See Eric Campbell and David Blumenthal, 'The Selfish Gene: Data Sharing and Withholding in Academic Genetics', *Science*, 31 May 2002

³ See Wesley Cohen et al, *Where Excludability Matters: Material v. Intellectual Property in Academic Biomedical Research* <http://siepr.stanford.edu/programs/SST_Seminars/walsh.pdf>, which illustrates the benefits of self-archiving.

This is why we brought in funders of disease research and institutional hosts of research from the beginning – this is the part of infrastructure that is social engineering, not software. The secrecy and competition do not maximise the likelihood of meaningful discovery coming from limited funding, and thus funders (especially of rare or orphan diseases) have a particular incentive to maximise the easy movement of biological materials to maximise follow-on research.

The MTA project covers transfers among non-profit institutions as well as between non-profit and for-profit institutions. It integrates existing standard agreements into a Web-deployed suite alongside new Science Commons contracts and allows for the emergence of a transaction system along the lines of Amazon or eBay by using the contracts as a tagging and discovery mechanism for materials.

This metadata driven approach is based on the success of the Creative Commons licensing integration into search engines and further allows for the integration of materials licensing directly into the research literature and databases so that scientists can ‘one-click’ inline as they perform typical research. And like Creative Commons licensing, we can leverage the existing Web technologies to track materials propagation and reuse, creating new data points for the impact of scientific research that are more dimensional than simple citation indices, tying specific materials to related peer-reviewed articles and data sets.

The MTA project was launched in collaboration with the Kauffman Foundation, the iBridge Network of university technology transfer offices, and neurodegenerative disease funders. It currently includes more than 5000 DNA plasmids covered under standard contracts and is available through the Neurocommons project described in the next section.

PROOF OF CONCEPT IN KNOWLEDGE SHARING: A SEMANTIC WEB FOR NEUROSCIENCE

In collaboration with the W3C Semantic Web Health Care and Life Science interest group, we are integrating information from a variety of standard sources to establish core interoperable content that can be used as a basis for bioinformatics applications. The combined whole is greater than the sum of its parts, since queries can cut across combinations of sources in arbitrary ways.

We are also providing an operational knowledge base that has a standard, open query endpoint accessible by Internet. The knowledge base incorporates information marshalled from more than a dozen databases, ontologies, and literature sources.

Entities discussed in the text, such as proteins and diseases, need to be specifically identified for computational use, as do the entities' relationships to the text and the text's assertions about the entities (for example, a particular asserted relationship between a protein and a disease). Manual annotation by an author, editor, or other 'curator' may capture the text's meaning accurately in a formal notation. However, automated natural language processing (including entity extraction and text mining) is likely to be the only practical method for opening up the literature for computational use.

We were only able to process the abstracts of the literature as the vast majority of the scientific literature is locked behind firewalls and under contracts that explicitly prevent using software to automatically index the full text where it is accessible. Although most papers run more than five pages, the abstracts typically were limited to a paragraph.

For tractability, we limited the scope to the organisms of greatest interest to health care and life sciences research: human, mouse, and rat. We are also providing the opportunity for interested parties to 'mirror' the knowledgebase and we encourage its wide reuse and distribution.

In combination with the data integration and text processing, we are also offering a set of analytic tools for use on experimental data. The application of prior knowledge to experimental data can lead to fresh insights. For example, a set of genes or proteins derived from high throughput experiments can be statistically scored against sets of related entities derived from the literature. Particular sets that score well may indicate what's going on in the experimental setting.

In order to help illustrate the value of semantic web practices, we are developing statistical applications that exploit information extracted from RDF data sources, including both conversions of structured information (such as Gene Ontology annotations) and relationships extracted from literature. The first tools we hope to roll out are activity centre analysis for gene array data and set scoring for profiling of arbitrary gene sets, donated to Science Commons by Millennium Pharmaceuticals.

Taken together, we call these three efforts the Neurocommons – an open source, open access knowledge management platform, with an initial therapeutic focus on the neurosciences. And we hope to use the Neurocommons both as a platform to facilitate knowledge sharing and to secure empirical evidence as to the value of shared knowledge in sciences.

CONCLUSION: CYBERINFRASTRUCTURE FOR KNOWLEDGE SHARING

The Neurocommons project is a very good start. It shows the potential of shared knowledge systems built on open content. And it has the potential to explode through horizontal downloading, editing, and reposting, as the Web exploded. The idea of connectivity via ‘viewing source’ is an explicit part of our design methodology, and our tools have already been picked up and integrated into such systems as the Mouse BIRN Atlasing Toolkit (MBAT), which was built from the combined efforts of groups within the Mouse BIRN (Biomedical Informatics Research Network, a distributed network of researchers with more than \$25 million in U.S. Government funding).

But the Neurocommons is, at root, a proof of concept. And from it we are learning some basic lessons about the need for infrastructure for knowledge sharing. Science Commons is on a daily basis forced to create namespaces, persistent URLs, and line after line of ‘plumbing code’ to wire together knowledge sources.

If we are going to get to the goal stated above, of dramatic increases in efficiency and radical transformation of outmoded discovery models, we are going to need a lot of infrastructure that doesn’t yet exist.

We need publishers to look for business models that aren’t based on locking up the full text, because the contents of the journals – the knowledge – is itself part of the infrastructure, and closed infrastructure doesn’t yield network effects. We need open, stable namespaces for scientific entities that we can use in programming and integrating databases on the open Web, because stable names are part of the infrastructure. We need real solutions about long-term preservation of data (long-term meaning a hundred years or more). We need new browsers and better text processing. We need a sense of what it means to ‘publish’ in a truly digital sense, in place of the digitisation of the

paper metaphor we have in the PDF format. We need infrastructure that makes it easy to share and integrate knowledge, not just publish it on the Web.

None of this is easy. Much of it is very, very hard. But the current system is simply not working. And the reward of pulling together what we already know into open view, in open formats, where geniuses can process and exploit it, could be a world in which it is faster, easier, and cheaper to find drugs and cure disease. This is possible. We just have to have the vision and courage to build the highways.

PART FOUR

SCHOLARLY
COMMUNICATIONS

WORKING FOR A RESEARCH– FRIENDLY IPR FRAMEWORK IN THE UK

Frederick Friend¹

INTRODUCTION

Research institutions and individual researchers in many countries are facing intellectual property issues which are changing the way in which the results of research are disseminated, how those results are used and by whom, and how current research feeds into future research. Some of the key questions which will be determined in part at least by intellectual property issues are:

- access: will the text and data in research papers be accessible and under what licensing conditions?
- publication: how will text and data be published, in journals or held in personal or institutional repositories?
- ownership: will authors, employers, funders or publishers claim ownership of text and/or data?
- re-use: will owners restrict re-use, even for academic purposes?
- management: how will text and data silos be managed and by whom?
- preservation: how will text and data be preserved and by whom?

These are key questions both for the current generation of researchers and also for future generations whose work may be helped or hindered – even prevented – by decisions being made now. The benefits flowing

¹ Scholarly Communication Consultant, JISC; Scholarly Communication UCL, Honorary Director. Email: f.friend@ucl.ac.uk.

from today's biomedical research would be impossible to achieve without the strong action taken by the research community a few years ago in opening the Human Genome Database for use without restriction. The commercial forces which almost locked away the genome data could lock away equally valuable research results in the future if the academic community does not ensure that appropriate intellectual property rights remain within the academic community. When commercial interests control rights in content generated within the academic sector, it is sometimes due to academic neglect of IPR issues. The work described in this chapter is informing the UK academic community of possibilities for the good management of research text and data.

THE UK JOINT INFORMATION SYSTEMS COMMITTEE

The UK Joint Information Systems Committee is addressing all of these issues as part of its role in providing guidance to UK universities and colleges on the changes taking place in the networked information environment. The JISC (Joint Information Systems Committee) is a sub-committee of the Higher Education Funding Council for England and the higher and further education funding councils in Scotland, Wales and Northern Ireland. JISC's activities support education and research by promoting innovation in new technologies and by the central support of ICT services. The JISC provides a world-class network (JANET, the Joint Academic Network), access to purchased and public-domain electronic resources, and projects to develop new environments for learning, teaching and research.

At almost every point in the JISC's activities, IPR issues have to be considered, and many of these issues are identical to or closely-related to the IPR issues UK universities are themselves facing in managing networked information. In order to find ways forward through some of the IPR issues UK universities and the JISC itself faces, a range of services and mechanisms have been developed. These include:

- JISC Legal: an information service for UK universities;²
- A dedicated IPR consultancy to provide expert advice to the JISC Development Group and its projects;³
- A specific Legal and Policy Cluster as part of the JISC Repositories Programme to encourage the sharing of IPR issues amongst projects;
- Collaboration with DEST in Australia and SURF in the Netherlands to share experience in exploring academic-related IPR issues.

Through these mechanisms the JISC is able to investigate the applicability of new IPR models such as Creative Commons and also design tools such as licence registries to support the rights management process. In parallel to this work the JISC Collections team negotiate academic-friendly licensing agreements as part of JISC's content activities, e.g. the right to make copies for preservation, using standardised rights expressions and model licences to help the use of content within the sector.

JISC IPR PROJECTS

The JISC funds many projects under various Programmes⁴ in order to assist UK universities in the introduction of new environments for learning, teaching and research. Virtually every project has to address IPR issues in respect of the project's own content and content that may be used if the project develops into a service. A project's experience in addressing IPR issues is invariably included as part of the project's final report, made available on the JISC web-site so that the UK academic community may benefit from the experience gained in the project.

In addition the JISC has commissioned a number of projects specifically to address IPR questions important in the academic sector. A selection of such specific IPR projects is:

² JISC, *JISC Legal* <<http://www.jisclegal.ac.uk/>>.

³ JISC, *IPR consultancy* <<http://www.jisc.ac.uk/whatwedo/projects/ipr/iprconsultancy.aspx>>.

⁴ A full list of JISC Programmes is available at <<http://www.jisc.ac.uk/whatwedo/programmes.aspx>>.

- The Intrallect DRM study completed in 2004, which revealed more interest in rights management from the teaching and learning community than from the research community;⁵
- The Rights and Rewards Project 2005–07, covering arrangements for repository deposit, balancing academic rewards from deposit with academic concern about loss of control over content;⁶
- The Copyright Knowledge Bank (part of ongoing JISC/SURF Partnering on Copyright), which aims to develop the SHERPA/RoMEO database as a tool for authors in retaining rights and for universities in developing repositories;⁷
- The JISC/SURF Licence to Publish (developed in 2006 and supported by ongoing advocacy), a Licence which is part of the JISC/SURF Copyright Toolbox and which was created to provide a model text authors to use instead of copyright assignment;⁸
- The Trust DR project, looking specifically at institutions' DRM systems for learning objects held in repositories.⁹

A common feature of the IPR services and projects funded by the JISC is that the JISC is placing an emphasis upon practical measures to effect cultural change. Respect for copyright is a strong feature of the JISC's policy, but some of the copyright structures suitable in a print environment do not fit well in a networked information environment. The adoption of new practices and structures may require cultural

⁵ Intrallect Ltd on behalf of JISC, *Digital Rights Management – Final Report* (2004) <<http://www.intrallect.com/drm-study/>>.

⁶ The Rights and Rewards Project <<http://rightsandrewards.lboro.ac.uk/>>.

⁷ JISC/SURF Partnering on Copyright, *Copyright Knowledge Bank* <<http://www.lboro.ac.uk/departments/lis/disresearch/poc/pages/knowledgebank.html>>.

⁸ JISC/SURF Copyright Toolbox, *Licence to publish* <<http://copyrighttoolbox.surf.nl/copyrighttoolbox/authors/licence/>>.

⁹ Trust DR <<http://trustdr.ulster.ac.uk/>>.

changes within the academic community in the community's attitude towards IPR issues.

JISC AND THE MANAGEMENT OF DATA

Many academic members of JISC Committees during the past year or two have mentioned the increasing importance of access to research data. The research information landscape is changing, and while text is still the common medium of communication in many subject areas, data and images are becoming of primary importance in other disciplines. The JISC has recognised this trend in its strategic documents.

The 'JISC Strategy 2007–2009'¹⁰ contains as Key Aim 3: 'To promote the development, uptake and effective use of ICT to support research',¹¹ and a Key Deliverable within this Key Aim is 'in collaboration with the Research Councils [to] provide a robust, trustworthy, secure, interoperable and scalable infrastructure for the transmission, storage, sharing, accessibility and dissemination of research data and outputs'.¹² This is a very strong strategic commitment to data provision. JISC recognises that access to and re-use of data is as important to the research community as access to and re-use of text. Legal issues are appearing regularly as JISC undertakes activities in fulfilment of this Key Deliverable.

As the JISC Executive looks to fulfil this strategic commitment, it is not surprising that many JISC-funded projects are about the management of data. A few examples are:

- The DCC SCARP Project is investigating different disciplinary approaches to data deposit, sharing and re-use, curation and preservation;¹³

¹⁰ JISC, *JISC Strategy 2007–2009* (2007) <http://www.jisc.ac.uk/media/documents/about_us/strategy/jisc_strategy_2007–2009.pdf>.

¹¹ JISC, *JISC Strategy 2007–2009* (2007) 15 <http://www.jisc.ac.uk/media/documents/about_us/strategy/jisc_strategy_2007–2009.pdf>.

¹² JISC, *JISC Strategy 2007–2009* (2007) <http://www.jisc.ac.uk/media/documents/about_us/strategy/jisc_strategy_2007–2009.pdf>.

¹³ DCC Scarp Project <www.dcc.ac.uk/scarp/>.

- The GRADE Project is investigating the technical and cultural issues around the re-use of geospatial data and facing up to real IPR issues;¹⁴
- The SPECTRa Project has been investigating arrangements for the deposit of chemistry data in repositories and amongst the findings is that 'IPR issues relating to the ownership and re-use of scientific data are complex';¹⁵
- The StORe Project has developed a demonstrator for linking source and output data repositories to enable the flow of data from workbench into publications;¹⁶
- The *International Study on the Impact of Copyright Law on Digital Preservation* is UK contribution to a collective study by four international organisations;
- The JISC IPR Consultancy has made available a paper on *IPR and Licensing issues in Derived Data* by Naomi Korn, Charles Oppenheim and Charles Duncan.¹⁷

Various aspects of data management are being explored by these and other projects. It is clear, for example, that any new arrangements for access to and re-use of data, even in an open access environment, will have to take into account differing disciplinary attitudes towards the sharing of data. A reality of the data situation is also that access to research data is not always within the control of the academic community and may involve partnership with commercial interests. In this complex environment researchers need guidance from their employers and funders on IPR issues in relation to data. In order that research can be conducted efficiently a workflow approach to data

¹⁴ See for example: GRADE, *Scoping a Geospatial Repository for Academic Deposit and Extraction* (2007) <<http://edina.ac.uk/projects/grade/gradeDigitalRightsIssues.pdf>>.

¹⁵ JISC, *Project SPECTRa: Submission, Preservation and Exposure of Chemistry Teaching and Research Data* (March 2007, Final Report) <<http://www.lib.cam.ac.uk/spectra/FinalReport.html>>.

¹⁶ StORe <<http://jiscstore.jot.com/WikiHome>>.

¹⁷ Naomi Korn, Charles Oppenheim and Charles Duncan, *IPR and Licensing issues in Derived Data* (2007) <<http://www.jisc.ac.uk/media/documents/projects/iprderiveddatareport.pdf>>.

access is needed but IPR conditions will need to follow the data as it travels through the process.

CONCLUSION

At present no clear answers are available to any of the questions set out in the Introduction to this chapter. The issues are complex, involve a mix of stakeholders with different priorities and concerns, and cannot be resolved by one country acting alone. Collaboration between stakeholders in various countries will be necessary in helping the research community find the answers it needs in order to work efficiently and cost-effectively. A research-friendly IPR framework will be essential to achieve this successful outcome. As part of this international effort the JISC is working to ensure research-friendly arrangements for access to and re-use of data and text produced by the UK research community.

CREATING A LEGAL FRAMEWORK FOR COPYRIGHT MANAGEMENT OF OPEN ACCESS WITHIN THE AUSTRALIAN ACADEMIC AND RESEARCH SECTOR¹

**Professor Brian Fitzgerald, Professor Anne
Fitzgerald, Professor Mark Perry, Scott Kiel-
Chisholm, Erin Driscoll, Dilan Thampapillai,
Jessica Coates**

INTRODUCTION

There is an increasing recognition, in Australia and internationally, that access to knowledge is a key driver of social, cultural and economic development. The argument for greater access to, and reuse of, research outputs is reinforced by the fact that much research in Australia is funded by public money and, consequently, that there is a public benefit to be served by allowing citizens to access the outputs they have funded.² This recognition poses both legal and policy challenges, in

¹ This chapter is derived from the OAK Law Project Report No. 1 ‘*Creating a Legal Framework for Copyright Management of Open Access Within the Australian Academic and Research Sector*’ by Professor Brian Fitzgerald (QUT), Dr Anne Fitzgerald (QUT), Professor Mark Perry (UWO), Scott Kiel-Chisholm (OAKLaw), Erin Driscoll (OAKLaw), Dilan Thampapillai (VU) and Jessica Coates (CCi). Special thanks to Kylie Pappalardo for her assistance.

² Markus Buchhorn and Paul McNamara, *Sustainability Issues for Australian Research Data: The Report of the Australian e-Research Sustainability Survey Project* (2006) Australian Partnership for Sustainable Repositories <<http://www.apsr.edu.au>> and <http://dspace.anu.edu.au/bitstream/1885/44304/1/aeres_report.pdf> at 23 October 2006 (hereinafter, Buchhorn and McNamara, *Issues for Research Data*). At page 26 the report states that in 2002–2003, 45 per cent of the \$12.25 billion expended on research and development in Australia was funded by government and in 2004, 90 per cent of the

terms of existing legal frameworks such as copyright law and traditional business models.

With the rise of networked digital technologies our knowledge landscape and innovation system is becoming more and more reliant on best practice copyright management strategies and there is a need to accommodate both the demands for open sharing of knowledge and traditional commercialisation models. As a result, new business models that support and promote open innovation are rapidly emerging.

This chapter analyses the copyright law framework needed to ensure open access to outputs of the Australian academic and research sector such as journal articles and theses. It overviews the new knowledge landscape, the principles of copyright law, the concept of open access to knowledge, the recently developed open content models of copyright licensing and the challenges faced in providing greater access to knowledge and research outputs.

THE NEW KNOWLEDGE LANDSCAPE

There have been fundamental changes in the framework within which knowledge is generated, accessed, disseminated and reused. The digital, networked environment and, in particular the widespread availability of broadband Internet access, is democratising creativity and innovation and has made it possible to process and construct knowledge in ways that were unimaginable only two years ago.

These changes have provided researchers and the general community with enormous possibilities for new forms of collaborative and serendipitous innovation. It is now in the hands of millions of people to readily produce and disseminate their own creative works; research groups can share information and develop collaborative synergies in ways that were not previously feasible.³ For example, blogs (Web logs), wikis, VoIP (Voice over Internet Protocol), podcasts and vodcasts are now commonplace, as are digital repositories.⁴ There has also been a

\$4.3 billion expended on research and development by higher education institutions was funded by government.

³ See Eric Von Hippel, *Democratizing Innovation* (2006) <<http://web.mit.edu/evhippel/www/democ.htm>> at 22 July 2006.

⁴ Neil Jacobs believes that technologies such as blogs, wikis and peer-to-peer repositories often come into universities and colleges ‘under the radar’. ‘The PROWE project

rise in collaborative projects such as Wikipedia – an online peer produced encyclopaedia also available on CD – now contains more than 4 million articles in 229 languages.

The legal challenges to this evolving landscape rest in the fact that while much of this research output can be presented at the click of a button it is often subject to copyright law and can only be used with permission of the copyright owner or on the basis of some other authorising principle or provision. The great challenge for this evolving knowledge landscape is, therefore, to build more efficient copyright ownership, management and licensing models that can be used to allow access to knowledge and prosper the research sector.

OVERVIEW OF PRINCIPLES OF COPYRIGHT LAW

Providing better access to research and knowledge through best practice copyright management can only be achieved by appreciating and understanding the scope and limitations of copyright law.

What is Copyright?

Copyright is a type of intellectual property founded on a person's creative skill and labour. It allows the copyright owner to control certain acts (such as copying) and to prevent others from using protected material without permission, unless an exception applies. A copyright owner has the right to take action for copyright infringement in the event that a person uses all, or a 'substantial part', of their copyright material in one of the ways exclusively controlled by the copyright owner, without their express or implied permission and where no defence or exception to infringement applies.

<<http://www.prowe.ac.uk>> is asking whether blogs and wikis in particular can be used to support the huge distributed networks of tutors associated with the Open and Leicester Universities. The SPIRE project <<http://spire.conted.ox.ac.uk/cgi-bin/trac.cgi>> is installing the secure Lionshare <<http://lionshare.its.psu.edu/main/>> peer-to-peer system, to explore its potential in teaching and learning and, in part, to dispel the mistaken notion that peer-to-peer equals Napster equals insecure and probably illegal activity.' Neil Jacobs, 'Digital Repositories in UK universities and colleges' (2006) Issue 200 *FreePint* 13–15 <<http://www.freepint.com/issues/160206.htm>> at 23 February 2006 (hereinafter Jacobs, *Digital Repositories*).

A person or organisation can also be liable for copyright infringement if they have authorised someone else to infringe copyright, to the extent that they sanction, approve or countenance the infringing conduct. For example, allowing PhD students to provide online access to a thesis knowing that the student has not obtained the prior permission of all the underlying rights holders (such as owners of copyright in pictures, graphics included in the thesis or accompanying audiovisual material) to digitise and communicate the work could potentially result in a university being held liable for authorising copyright infringement.

What Type of Material Does Copyright Protect?

For copyright to subsist material must fall within a category recognised under the *Copyright Act 1968* (Cth) (*Copyright Act*): namely, original literary, dramatic, musical and artistic works, as well as sound recordings, films, sound and television broadcast and published editions. Therefore, copyright protects not only written material (such as books, theses and reports) and creative works such as photographs, paintings and multimedia works but also scientific and technical creations (for example computer software and datasets).

What Rights Does Copyright Protect?

The exact nature of the rights granted to copyright owners will depend on the nature of the material being protected. However, in general they will include the exclusive right of reproduction, publication, performance, communication and adaptation. As with all IP rights, the exclusive rights provided by copyright are intangible in nature, generally granted for a limited time (for example, either 70 years from the death of the creator of a work or 70 years from first publication of a film or sound recording), and are distinct from the physical property in which protected material is embodied.

Balancing the Interests of Copyright Owners and Users – Exceptions to Copyright

Most copyright laws have been structured to provide a balance between the provision of incentives in the area of innovation and creativity and achieving the public interest goal of encouraging education, research, the

free flow of information and freedom of expression,⁵ while also being careful not to restrict competition in the marketplace. To give effect to this balance, the *Copyright Act* contains a range of ‘free use’ or ‘blanket’ exceptions to copyright infringement which allow material to be used without the permission of, or a licence from, the copyright owner, together with a range of statutory licences which allow the making and communication of multiple reproductions of certain works for set licence fee, thereby reducing overall administration and transactional costs.

The copyright exceptions of relevance to the education and research sector include the fair dealing exceptions for research and study, criticism and review and reporting the news. These exceptions are necessarily limited in that the dealing must have been performed for one of these four purposes, and it must be considered to be ‘fair’. Thus, there is no open defence such as general fair dealing or fair use under Australian copyright law.

The statutory licences for the education sector enable educational institutions to copy television and radio programs off-air and to reproduce and communicate print copyright works and electronic versions of literary, dramatic, artistic and musical works for educational purposes, in return for payment to declared collecting societies.

Rights Related to Copyright — Moral Rights and Performers’ Rights

In addition to the traditional economic rights discussed above, the *Copyright Act* also bestows certain moral rights and performers’ rights.

Australian copyright law grants performers both economic⁶ and personal rights over audio (but not audiovisual) recordings of their performances. These rights consist of:

⁵ Copyright Law Review Committee, *Copyright and Contract* (2002) 24 <www.clrc.gov.au> at 22 July 2006.

⁶ The economic rights for performers in sound recordings became effective from 1 January 2005. Section 22(3A) of the Act provides that the performer and the owner of any sound recording of the performance own the copyright jointly, subject to any agreement to the contrary. Commissioned sound recordings for which the performer is paid a fee, or those made under an employment contract, are owned by the commissioner or employer (section 97(3)).

- the right to authorise the recording and communication of live performances (and distributions of recordings of live performances);⁷
- copyright in sound recordings;⁸ and
- moral rights in performances.⁹

The first two of these rights only apply to performances that took place after 1 October 1989. A performer's rights to authorise recording and communication of their performances or the reproduction or performance of recordings last for 50 years from the date of the performance. Rights to authorise communication of recordings or the use of a recording in a soundtrack last for 20 years from the date of recording.¹⁰

Individual creators of literary, dramatic, musical and artistic works and films¹¹ have the following moral rights in relation to works or films they have created: the right to be attributed (credited) for their work, the right not to have their work falsely attributed and the right not to have their work treated in a derogatory way.

Performers also have moral rights in relation to live performance, so far as the performance consists of sounds, or a sound recording of a live performance. These rights apply to live performances as defined in the *Copyright Act* which include expressions of folklore and musical, dramatic and dance performances.¹² The moral rights granted to performers mirror the moral rights in traditional works. Generally, they will last for the duration of the copyright in the sound recording, although the right of integrity in a recorded performance only lasts until the performer's

⁷ See *Copyright Act 1968* (Cth), pt XIA.

⁸ *Copyright Act 1968* (Cth) s22(3A). This right is subject to any agreement to the contrary, and does not apply to commissioned performances or performances conducted in the course of employment - s 97(3).

⁹ *Copyright Act 1968* (Cth), ss 195AXA, 195AXB and 195AXC.

¹⁰ *Copyright Act* sections 248CA(3), 248G(1) and (2)).

¹¹ In relation to a film, the director, producer and screenwriter all separately own moral rights in relation to a film and where there are multiple directors etc. it is only the principal director, screenwriter and producer who hold moral rights.

¹² *Copyright Act 1968* (Cth) s 22 (7).

death.¹³ The same reasonableness exemptions that apply to traditional moral rights also apply to performers' moral rights.¹⁴ Furthermore, to make the authorisation process efficient for performances involving multiple performers, the Copyright Act permits an agent acting for a group of performers to grant permission to reproduce etc any sound recordings.¹⁵

Technological Protection of Copyright Material

Digital technology has made it possible to easily reproduce and communicate copyright material in near perfect form. Copyright owners have, therefore, sought — as an alternative to traditional forms of legal protection — to rely on technology to prevent others from using their work without their permission. However, the *Copyright Act* also provides legal recognition for new mechanisms for copyright owners to protect and enforce their rights. For example, the *Copyright Act* provides legal protection for the use of electronic rights management information (ERMI) (such as digital watermarks) to described, identify, monitor and track digital copyright material. These rights, in effect, potentially enable a copyright owner to monitor every access and use of their copyright material.

The *Copyright Act* also contains specific provisions which reinforce the use of technology, in the form of digital locks (known as technological protection measures (TPMs)) to regulate access and further copying of copyright material. It is a civil infringement and/or a criminal offence (the level of liability depending on the circumstances of the infringement) under sections 116AO and 116AP and 132APD and 132APE of the *Copyright Act* to deal in circumvention devices or services, including the manufacturing, importing, distribution (including online), provision and offering to the public of circumvention devices and services. Under section 116AN(1) of the *Copyright Act*, a copyright owner or exclusive licensee of copyright in a work or other subject matter may bring an action against a person who does an act resulting in the circumvention of a TPM protecting the work or other subject

¹³ *Copyright Act 1968* (Cth), s 195ANA.

¹⁴ *Copyright Act 1968* (Cth), ss 195AXD and 195AXE.

¹⁵ See *Copyright Act 1968* (Cth), ss 113A and 191B.

matter, where that person knows or ought reasonably to know that their act would result in circumvention of the TPM. Criminal penalties may also apply under section 132APC(1) where the circumvention was done with the intention of obtaining a commercial advantage or profit. The *Copyright Act* contains a set of exceptions which allow the circumvention of TPMs for certain permitted purposes (such as security testing or error correction).

Copyright Licensing

Despite legal recognition of copyright owners' rights to embrace technology to better control access to, use and dissemination of copyright material in the digital environment, general principles of copyright law, through mechanisms such as licensing, can also support open access to knowledge.

While it is possible to either sell or give away copyright via either an assignment, transfer or through a bequeath, it is equally possible for copyright owners to share copyright between themselves and third parties under a licence. A licence is a 'permission' or form of authorisation from the copyright owner to use the copyright material in one or more of the ways which falls within the copyright owner's exclusive rights. A licence can be exclusive, non-exclusive or implied.

Under an exclusive licence the licensee (in other words, the recipient of the licence) is the only person who can use the works in the way or ways covered by the licence (even to the exclusion of the copyright owner). A non-exclusive licence merely provides a user/third party with the right to exercise one or more of the copyright owner's rights in the work but not to the exclusion of the copyright owner or other licensees. Therefore, a copyright owner may grant multiple and simultaneous non-exclusive licences.

It is also important to note that with both assignments and licences, copyright can be divided in a number of ways, including by territory, time and type of use. For example, a licence can give a person permission to reproduce a work, without giving permission to publish or communicate the work. Similarly, a licence may give a publisher the right to publish the material only in Australia, or only until a certain date. The various licensing models for managing access to research findings are outlined in 'Open Access and Open Content Licensing' below.

THE LEGAL PROTECTION OF DATABASES – A SPECIAL CASE

Open access can be pursued not only in relation to academic and research output in traditional forms (such as, research proposals, project plans, summaries of research results, conference papers, journal articles and books in published form) but also in relation to new forms of output such as data files, complex databases involving compilations of datasets and embedded software and multimedia works.

In developing systems designed to promote open access to knowledge in the Australian academic and research sector, and to data in particular, academics and researchers need to consider:

- the copyright status of the database and whether the data is protected by copyright
- whether third party copyright is affected by making a database available to the public
- the type of legal or technological measures that can be used to protect a database.

Whether Databases are Protected by Copyright

As a general principle, copyright law protects the expression of an idea and not the idea itself. To this end data, without more, is not protected by copyright law. The compilation of data, however, is protected to varying degrees by copyright law in different jurisdictions throughout the world. In the United States (US) and the European Union (EU) data compilation — selection and arrangement of the data — is protected where there is an element of intellectual creation. In addition to copyright protection available for databases, Europe also has a *sui generis* database right which may protect non-original databases that do not attract copyright protection but which are nevertheless valuable and have required substantial economic investment.¹⁶

In Australia, databases may attract copyright protection if the creation of the database has involved sufficient expenditure of time, money, skill or

¹⁶ On Europe, see the European Union Database Directive, available at <http://ec.europa.eu/internal_market/copyright/prot-databases/prot-databases_en.htm> at 8 July 2008.

effort to satisfy the threshold level of originality required in order for copyright to subsist in a literary work. In the recent case of *Desktop Marketing Systems Pty Ltd v Telstra Corporation Limited*¹⁷ (*Desktop Marketing*) the Full Court of the Federal Court held that the mere arrangement of names in alphabetical order in a phone book was sufficient to found copyright protection. As a result, the standard of originality for copyright protection in Australia is considerably lower than other jurisdictions.¹⁸ Thus, it is the case that where facts are compiled through industrious labour (in other words where the intellectual effort is very low or non-existent) they will receive a higher degree of protection in Australia than in other jurisdictions.

In addition to the broad scope of protection available for databases in Australia, the very narrowly defined nature of the fair dealing exceptions (as explained ‘Overview of Principles of Copyright Law’ above) confers further control for owners of copyright in databases.

Practical Measures for Database Compilers to Protect Their Copyright

From a practical standpoint, database compilers need to identify the uses of their database that they wish to allow. They then need to put in place the relevant agreements to facilitate those uses. This involves identifying and, where necessary, obtaining copyright permissions from third party copyright owners. It also involves preparing agreements that clearly set out the conditions of use of the database. In addition database owners could employ (TPMs) to regulate the use of a database, or they could seek to adopt a range of licensing models such as open content licensing like Creative Commons licences.

Third Party Content

When researchers develop databases containing information from a range of sources, copyright in some of the materials selected for inclusion will belong to third parties (in other words, commercial publishers, governments, individual authors and research institutes).

¹⁷ [2002] FCAFC 112.

¹⁸ See also *Nine Network Australia Pty Ltd v IceTV Pty Ltd* [2008] FCAFC 71.

However, when the researcher makes the database available for access by other researchers, it will be necessary to ensure that the researcher has the legal authority to do so, either under a recognised exception or through a licence.

Where a licence is used to obtain permissions by third party owners of copyright material included in the database, the licence should sufficiently permit the researcher to authorise other persons to use the material in the way in which the database compiler and database users wish to use the material. If the licence does not do so, release of copyright material owned by third parties will infringe their copyright.

OPEN ACCESS

With the growth of the new digital and virtual knowledge landscape, we have seen the potential for greater control over access and usage by copyright owners. The rising costs of subscriptions to key academic journals, in large part made possible by, and implemented through, the first generation of digital distribution and licensing models, has motivated a frustrated research community into finding new ways to disseminate knowledge. Faced with the enormous potential of the Internet and the increasing limitations presented by traditional journal licensing, researchers worldwide have united in the Open Access (OA) movement which aims to disseminate knowledge broadly and freely across the Internet in a timely fashion (especially that which is publicly funded).¹⁹ User lead movements such as Open Access (OA) and Free/Libre and Open Source Software (FLOSS)²⁰ have sought to utilise the great advances in information and communication technologies to

¹⁹ In 1991, the first free scientific online archive, arXiv, was created at Los Alamos but it is now hosted by Cornell University. The fields covered include physics, mathematics, non-linear science, computer science and quantitative biology. See <<http://www.lib.mtu.edu/eresources/eresearch/searchresults.aspx?publisherid=240>> and <<http://arxiv.org/>>.

²⁰ See further: Glyn Moody, *Rebel Code: Linux and the Open Source Revolution* (2001). See also: Lawrence Lessig, *The Future of Ideas: The Fate of the Commons in a Connected World* (2001) 50 ff; Sam Williams, *Free as in Freedom: Richard Stallman's Crusade for Free Software* (2002); Eric Raymond, *The Cathedral and the Bazaar* <<http://www.catb.org/~esr/writings/cathedral-bazaar>> at 22 July 2006; Brian Fitzgerald and Nic Suzor, 'Legal Issues For the Use of Free and Open Source Software in Government' (2005) 29 *Melbourne University Law Review* 412.

make research outputs more easily and immediately accessible and to promote a collaborative and participatory knowledge paradigm. This has resulted in the development of a worldwide network of institutional and disciplinary repositories containing numerous research outputs which use advanced Internet computing and Grid technologies to enable direct and shared collaboration amongst researchers in the form of e-Research. In Australia there are initiatives like E Print and Digital Theses Repositories and large supercomputing projects based around bio-informatics and geo-spatial data.²¹

Open Access Movement

Core Principles of OA

The core principle of OA is to open up access to research and scholarship, especially that which is publicly funded.²² This principle has been endorsed and further developed in the following declarations: Budapest Open Access Initiative (2002) (BOAI),²³ the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (2003) (Berlin Declaration),²⁴ and the Bethesda Statement on Open Access Publishing (2003) (Bethesda).²⁵

The Berlin Declaration's definition of Open Access contribution mirrors the definitions drafted in the BOAI and Bethesda Statement:

There are three main essentials: free accessibility, further distribution, and proper archiving:

Open access is real open access if:

²¹ The Australian Partnership for Advanced Computing (APAC) has been a key player in building this framework over the last six years <<http://www.apac.edu.au/>> at 31 July 2006.

²² Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (2003), <<http://www.zim.mpg.de/openaccess-berlin/berlindeclaration.html>>, and the Bethesda Statement on Open Access Publishing (2003), <<http://www.earlham.edu/~peters/fos/bethesda.htm>>.

²³ See <<http://www.soros.org/openaccess/read>>.

²⁴ See <<http://www.zim.mpg.de/openaccess-berlin/berlindeclaration.html>>.

²⁵ See <<http://www.earlham.edu/~peters/fos/bethesda.htm>>.

1. The article is universally and freely accessible, at no cost to the reader, via the Internet or otherwise, without embargo
2. The author or copyright owner irrevocably grants to any third party, in advance and in perpetuity, the right to use, copy, or disseminate the article, provided that correct citation details are given
3. The article is deposited, immediately, in full and in a suitable electronic form, in at least one widely and internationally recognized open access repository committed to open access and long-term preservation for posterity.²⁶

Another significant document representing a major international step forward in promoting open access to knowledge, and more broadly the sharing of knowledge, is the Draft Treaty on Access to Knowledge (A2K Treaty).²⁷

The A2K Treaty is largely a result of the work of Brazil and Argentina who, in August 2004, discussed a possible treaty concerning access to knowledge as part of the development agenda for the World Intellectual Property Organisation.²⁸ Amongst many purposes and objectives, the A2K Treaty is seeking to enhance the sharing of the benefits of scientific advancement and to promote new incentives to create and share knowledge resources without restrictions on access.²⁹ Article 1–1 of the A2K Treaty provides that the main objectives of the treaty are to protect and enhance access to knowledge, and to facilitate the transfer of technology to developing countries. Key areas which the A2K Treaty covers are: provisions regarding limitations and exceptions to copyright and related rights; patents; expanding and enhancing the knowledge commons; the promotion of open standards; the control of anticompetitive practices; authors' and performers' rights; and the transfer of technology to developing countries.

²⁶ Open Society Institute, *Open Access Publishing and Scholarly Societies A Guide* (2005) 6 <www.soros.org/openaccess/scholarly_guide.shtml> at 19 December 2005.

²⁷ See <<http://www.access2knowledge.org/cs/a2k>>.

²⁸ See <<http://www.access2knowledge.org/cs/a2k>>.

²⁹ In the preamble.

Access to Knowledge as a Human Right

The principle of open access to knowledge can also find a legal basis in international human rights laws, some of which clearly provide that people should have the right to hold private property, including IP rights. For example, the clearest enunciation of the right to hold private property is found in Article (Art) 27(2) of the *Universal Declaration of Human Rights* (UDHR).³⁰

However, this obligation is not absolute and must be read in the context of international human rights law that supports access to knowledge; for example:

- Art 17 of the *Convention on the Rights of a Child*³¹
- Art 19 of the *Universal Declaration of Human Rights*³²
- Art 13 of the *International Covenant on Economic, Social and Cultural Rights*³³
- Art 1.1 of the *International Convention on Cultural and Political Rights*.³⁴

International declarations, conventions and covenants are important in that they may also act as an interpretative guide when courts are called on to define the ambit of IP rights.³⁵

³⁰ Art 27(2) provides: '[e]veryone has the right to the protection of the moral and material interests resulting from any scientific, literary or artistic production of which he is the author.' <<http://www.unhchr.ch/udhr/lang/eng.htm>>

³¹ See <<http://www.ohchr.org/english/law/pdf/crc.pdf>>.

³² See <<http://www.unhchr.ch/udhr/lang/eng.htm>>.

³³ See <http://www.unhchr.ch/html/menu3/b/a_ceschr.htm>.

³⁴ See <<http://www.ohchr.org/english/law/ccpr.htm>>.

³⁵ *Minister of State for Immigration and Ethnic Affairs v Ab Hin Teob* (1995) 183 CLR 273, [25]-[26] <<http://www.austlii.edu.au/au/cases/cth/HCA/1995/20.html>>. As at 27 June 2006, this case has been applied in 34 subsequent decisions. See also *Kruger v Commonwealth* (1997) 190 CLR 1; *Horta v Commonwealth* (1994) 181 CLR 183; *Newcrest Mining (WA) Ltd v Commonwealth* (1997) 190 CLR 513; *Kartinyeri v Commonwealth* (1998) 195 CLR 337. See also Bryan Horrigan and Brian Fitzgerald, 'International and Transnational Influences on Law and Policy Affecting Government' in Bryan Horrigan (ed), *Government Law and Policy: Commercial Aspects* (1998) 2; Brian Fitzgerald, 'International Human Rights and the High Court of Australia' (1994) 1 *James Cook University Law Review* 78.

Key Features of Open Access

Peter Suber of Earlham College, states that the OA movement:

- proposes that authors electronically publish (or ‘archive’) pre-prints of their papers, in a manner analogous to Departmental Working Papers series of bygone days
- recommends the establishment of ePrints Archives by universities and other research institutions (to provide a manageably small number of persistent, professionally-managed and readily-discoverable locations, rather than tens of thousands of ephemeral, personal web-sites)
- publishes software that enables such ePrints archives to be managed
- recommends use of the Open Archive Initiative metadata standard, in order to support cross-discovery services
- approaches journal publishers to sanction author self-archival (already with great success)
- communicates with governments, with a view to ensuring that government policy and amendments to copyright law support and not undermine open access to authors’ pre-prints.³⁶

Support for Open Access

There has been significant support for the OA movement at the international level. As at 8 July 2008, 250 organisations around the world have signed the Berlin Declaration.³⁷ At the local level, various organisations have endorsed the principles of OA through developing organisation-specific declarations or policies on the topic. For example, some tertiary institutions recommend (or even mandate) that staff deposit their papers in the institutional repositories and many tertiary

³⁶ Roger Clarke, ‘A proposal for an open content licence for research paper (Pr)ePrint’ (August 2005) 10(8) *First Monday* <http://firstmonday.org/issues/issue10_8/clarke/index.html> at 22 November 2005.

³⁷ <<http://oa.mpg.de/openaccess-berlin/signatories.html>> at 8 July 2008..

institutions make the submission of post-graduate research papers and PhD theses into the institutional repository mandatory. For example, the world's two largest funders of medical researchers, the United Kingdom's Wellcome Trust³⁸ and the United States National Institutes of Health,³⁹ adopted, in 2005 and 2008 respectively, policies with a requirement to provide OA to the results of successful grantees. Such support of OA arguably benefits society by enabling access to medical research that can be used to save lives or enhance the quality of life.

NEW LICENSING MODELS

One of the most significant responses to the technological advances that have revolutionized the creation and distribution of copyright materials during the last decade has been the development of new systems for licensing (or authorising) others to obtain access to, and make use of, the protected material. These new forms of licences — usually referred to as 'open content' — are founded upon an acknowledgement of the existence of copyright in materials embodying knowledge and information. As mentioned in 'Overview of Principles of Copyright Law' above, copyright law makes it unlawful to reproduce and communicate copyright material unless the permission of the copyright owner or some other form of authorisation has been obtained. Therefore, while I might place an article in an institutional repository, if I say nothing more, the 'all rights reserved' default position will most likely apply, meaning that the end user's rights to engage in reproduction or communication of the material as an act of reuse will be unclear. Users may be able to read it online or print a copy but can they post an enhanced version on another website or make 30 copies for their students in class? In order to deal with these questions and to provide greater legal certainty and fluidity to the act of sharing knowledge, we have seen the rise of open content licensing (OCL). Running with the copyright material to which they are attached, open content licences identify materials that are available for reuse and grant permissive rights to users, thereby facilitating access and dissemination. In comparison to licences commonly used before the advent of the digital era they are

³⁸ See <<http://www.wellcome.ac.uk/>> and <http://en.wikipedia.org/wiki/Wellcome_Trust>.

³⁹ See <<http://www.nih.gov>> at 6 May 2006.

standardised, conceptually interoperable with other open content licences, machine (computer) enabled and they eliminate (or at least minimise) transaction costs (as they are automated). The best known of the open content licensing systems are those developed by the Creative Commons (CC) project and its associated Science Commons project.⁴⁰ Other projects employing open content licensing models include AEShareNet,⁴¹ a collaborative system designed to streamline copyright licensing to enable the more efficient development, sharing and adaptation of Australian educational materials.⁴²

Open Content Licensing

Open Content Licences are essentially voluntary IP licensing agreements, designed to provide an effective model for managing copyright in digital content. These agreements call on IP owners to consider sharing knowledge with the world through a legal mechanism that will allow a broad ambit of reuse. While OA aims to have research disseminated rapidly through the Internet, OCL aims to ensure that downstream user rights are clear. Therefore, OCL is not anti-copyright — it uses copyright as the basis for structuring open access.

A range of open content licences exist including:

- Creative Commons Licenses⁴³
- AEShareNet Instant Licences⁴⁴
- Design Science License⁴⁵
- GNU Free Documentation License⁴⁶
- Open Content License⁴⁷

⁴⁰ See <<http://creativecommons.org>> and <<http://sciencecommons.org>>.

⁴¹ <<http://www.aesharenet.com.au/>> at 22 July 2006.

⁴² See further Intrallect Ltd (Ed Barker, Charles Duncan) and AHRC Research Centre (Andres Guadamuz, Jordan Hatcher and Charlotte Waelde) *Final Report to the Common Information Environment Members of a study on the applicability of Creative Commons Licenses* (2005) Ch 3.6 <<http://www.intrallect.com/cie-study/>> at 22 July 2006.

⁴³ See <<http://creativecommons.org>>.

⁴⁴ See <<http://www.aesharenet.com.au/coreBusiness/#Instant>>.

⁴⁵ See <http://en.wikipedia.org/wiki/Design_Science_License>.

⁴⁶ See <http://en.wikipedia.org/wiki/GNU_Free_Documentation_License>.

- Open Directory Project License⁴⁸ used by Open Directory Project⁴⁹
- Open Game License⁵⁰ – Licence of the Open Gaming Foundation,⁵¹ as drafted by Wizards of the Coast⁵²
- Open Publication License⁵³ – Licence for the Open Content Project⁵⁴
- Open Game License⁵⁵ – Licence of the Open Gaming Foundation,⁵⁶ as drafted by Wizards of the Coast⁵⁷
- Open Publication License⁵⁸ – Licence for the Open Content Project⁵⁹
- The Commonwealth of Australia, represented by the Office of Spatial Data Management (OSDM), Spatial Data Licence used by Geoscience Australia.⁶⁰

As well as providing an effective model for managing copyright in digital content, OCL also has the added benefit of making copyright more active in the sense of enabling copyright material left inactive in archives (such as, government, public film or television authorities, museums) to be ‘licensed out’ and reused in an inexpensive and generic manner.

⁴⁷ See <http://en.wikipedia.org/wiki/Open_Content_License> and <<http://opencontent.org/opl.shtml>> at 6 May 2006.

⁴⁸ See <http://en.wikipedia.org/wiki/Open_Directory_Project_License>.

⁴⁹ See <http://en.wikipedia.org/wiki/Open_Directory_Project>.

⁵⁰ See <http://en.wikipedia.org/wiki/Open_Game_License>.

⁵¹ See <http://en.wikipedia.org/wiki/Open_Gaming_Foundation>.

⁵² See <http://en.wikipedia.org/wiki/Wizards_of_the_Coast>.

⁵³ See <http://en.wikipedia.org/wiki/Open_Publication_License>.

⁵⁴ See <http://en.wikipedia.org/wiki/Open_Content_Project>.

⁵⁵ See <http://en.wikipedia.org/wiki/Open_Game_License>.

⁵⁶ See <http://en.wikipedia.org/wiki/Open_Gaming_Foundation>.

⁵⁷ See <http://en.wikipedia.org/wiki/Wizards_of_the_Coast>.

⁵⁸ See <http://en.wikipedia.org/wiki/Open_Publication_License>.

⁵⁹ See <http://en.wikipedia.org/wiki/Open_Content_Project>.

⁶⁰ See <https://www.osdm.gov.au/osdm/docs/internet_licence.htm> and <https://www.ga.gov.au/products/servlet/controller?catno=63643&catno=63680&catno=63695&catno=63722&event=FILE_SELECTION>.

Open Content Licences can be also be seen to promote sustainable business models as they commonly adopt a dual licensing approach — in the sense that an OCL provides open access for non-commercial purposes but restricts reuse for commercial purposes. For example, the CC licences referred to above, provide that anyone can use the content subject to providing *attribution* to the author (BY) and any one or a number of the following optional conditions:⁶¹

- *non-commercial* distribution (NC)
- that *no derivative* materials based on the licensed material are made (in other words, all copies are verbatim) (ND)
- *share and share alike* (others may distribute derivative materials based on the licensed material under a licence identical to that which covers the licensed material) (SA).

Therefore, if a person writes an article on the legal aspects of downloading MP3s off the Internet, they might put that up on their website with an OCL licence such as a Creative Commons licence allowing the user to reproduce, recast and communicate the content so long as they provide attribution (BY), do not use it for a commercial purpose (NC) and share their innovations with the people of the world (SA). Thus a person can give permission in advance for their content to be used for non-commercial purposes before it can be used commercially.

CREATING LEGAL FRAMEWORKS FOR OPEN ACCESS TO ACADEMIC AND RESEARCH MATERIALS

As discussed throughout this chapter, there is increasing interest in ensuring that the output of publicly funded academic and research work is accessible and widely disseminated through open access channels.

It is essential to appreciate at the outset that, from the legal perspective, it is not possible to establish any kind of open access system simply by default. Rather, development of an open access system can only

⁶¹ Note that the ND and SA terms are mutually exclusive.

successfully occur through deliberate construction and active management.⁶²

In establishing the legal framework for a system of open access to academic and research materials, it is necessary for the key institutional players to:

- determine the degree of ‘openness’ required in relation to those materials⁶³
- understand the roles of, and relationships among, the relevant parties involved in funding, creating, publishing, distributing and using academic and research materials
- consider how best to manage the often complex inter-relationships among the various parties, especially with respect to their copyright interests in the materials, so

⁶² This point is reflected in Principle 1 of the Zwolle Principles (2003) which states:

1. Achievement of [the overall objective] requires the optimal management of copyright in scholarly works to secure clear allocation of rights that balance the interests of all stakeholders.

See

<<http://www.surf.nl/copyright/keyissues/scholarlycommunication/principles.php#Principles>> at 16 July 2006.

⁶³ There are various statements/declarations on open access in the context of academic materials, including: The Bethesda Statement on Open Access Publishing (2003)

<<http://www.earlham.edu/~peters/fos/bethesda.htm>> at 16 July 2006; the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (2003)

<<http://www.zim.mpg.de/openaccess-berlin/berlindeclaration.html>> at 16 July 2006; the Zwolle Principles (2003)

<<http://www.surf.nl/copyright/keyissues/scholarlycommunication/principles.php>> at 16 July 2006; the Budapest Open Access Initiative (2002)

<<http://www.soros.org/openaccess/view.cfm>> at 14 July 2006; and the Bermuda Principles (1996) <http://en.wikipedia.org/wiki/Bermuda_Principles> at 14 July 2006.

In addition to the numerous articles and blogs dealing with open access, there is now an emerging literature. Recently published books include: Richard Jones, Theo Andrew and John MacColl, *The Institutional Repository* (2006); N Jacobs (ed), *Open Access: Key Strategic, Technical and Economic Aspects* (2006) (most of the chapters are self-archived at

<http://www.earlham.edu/~peters/fos/2006_07_16_fosblogarchive.html#115325936391251995>, accessed 30 July 2006); John Willinsky, *The Access Principle: The Case for Open*

Access to Research and Scholarship (2005) available in part at

<<http://mitpress.mit.edu/catalog/item/default.asp?tid=10611&tttype=2>> at 30 July 2006.

that the relationships and copyright interests can be effectively managed to achieve the desired degree of open access in the system.

Developing and Publishing a Policy on Open Access

Before implementing a copyright management policy for the provision of access to and reuse of research outputs, each institution should develop and publish its policy on open access, clearly enunciating its objectives and interests in providing materials by this means.⁶⁴ This involves clearly identifying, articulating and observing the following:

- the categories of materials that are to be made available by open access
- the scope of open access which is to be afforded, in terms of the classes of persons who are to be allowed access and the extent of rights granted to access and re-use the materials.

Each institution should formally allocate responsibility to a suitability experienced and resourced office within the institution for implementation of the OA policy and for periodically reviewing its operation.

Mapping the Network of Legal Relationships

To ascertain who is permitted to use academic materials (in other words, that are available in a repository) and the extent of the permitted use of such materials, it is necessary to identify the various stakeholders and their respective roles, describe the legal relationships among them and understand how copyright interests are allocated.⁶⁵

⁶⁴ For example, the Zwolle Principles (2003) state the overall Objective as follows:

To assist stakeholders—including authors, publishers, librarians, universities and the public—to achieve maximum access to scholarship without compromising quality or academic freedom and without denying aspects of costs and rewards involved.

See: <<http://www.surf.nl/copyright/keyissues/scholarlycommunication/principles.php#Principles>> at 16 July 2006.

⁶⁵ This point is reflected in Principle 5 of the Zwolle Principles which states: ‘Copyright Management should strive to respect the interests of all stakeholders involved in the use

To date, much of the literature and research on copyright issues in open access systems has failed to adopt a sufficiently broad perspective which encompasses not only the full range of stakeholders involved but also the way the legal relationships among them impact upon the rights of repositories and end users. In particular, in considering rights to use materials deposited in repositories, much of the discussion has been overly focused on the Author-Publisher relationship, as defined in the Publishing Agreement (see below). Further, this already narrow focus has been channelled even more narrowly by the fact that much of the discussion has considered only those situations where copyright is assigned (or exclusively licensed) by the copyright owner (usually the author) to the publisher. The broader range of possible arrangements in relation to copyright ownership — including retention of copyright by the author — has received insufficient attention. To fail to adopt a broader perspective on the relationships between all the relevant stakeholders means a loss of the opportunity to achieve the most efficient and effective open access system by leveraging all the factors that can be brought to bear in pursuit of the open access objective.

The key stakeholders and relationships that will come into play in the structuring of an open access system are:

- A. **Funding organisation – Author:** the relationship between the organisation providing grants of funding for research and the author of outputs (such as, academic articles and research reports) of the funded research project, or the author's university or research institution [*Funding Agreement*].
- B. **Author – Employer:** the relationship between the author of academic or research output and their employer (such as, university or research institution) [*Employment Agreement and IP Policy*].
- C. **Author – Publisher:** the relationship between the author (or another party who owns copyright in works produced by the author, such as, the author's employer) and the publisher [*Publishing Agreement*].

and management of scholarly works; those interests may at times diverge, but will in many cases coincide.' See:

<<http://www.surf.nl/copyright/keyissues/scholarlycommunication/principles.php#Principles>> at 16 July 2006.

- D. **Author – Digital repository:** the relationship between the author (or another party who owns copyright in the author's works, such as, the author's employer or the publisher) and the digital repository in which a copy of the author's article is deposited [*Repository Deposit Licence*].
- E. **Digital repository – End users:** the relationship between the digital repository in which the author's article is deposited and persons who are authorised to access it (which may be the public at large or may be restricted to a particular group with defined access rights) [*Repository Distribution (End User) Agreement*].
- F. **Author/Publisher – End users:** the relationship between the author/publisher (or other owner of copyright, such as, the author's employer) and end users (in other words, persons authorised to access and use the material) [*Distribution Agreement*].
- G. **Copyright Collecting Society – Digital Repository and End Users:** Much of the administration of copyright in the educational context in Australia occurs pursuant to statutory licences administered by copyright collecting societies such as the Copyright Agency Limited (CAL), which collect fees from educational institutions as compensation for educational use of copyright materials. In establishing a system to enable access to academic and research materials in online repositories, it is necessary to consider how such materials will be treated under the statutory licence for reproduction and communication of works in electronic form under Division 2A of Part VB (ss 135ZMA to 135ZME of the *Copyright Act* [*Educational Statutory Licence*].

Each of these relationships and the particular copyright management issues they raise are considered, in turn, below.

A. Funding Organisation – Author / Research Institution (Funding Agreement)

Where research is being funded by an external source, that organisation may impose conditions on the researcher or recipient institution in relation to how the output of the funded research will be made available.

For example, it would not be unusual for a funding organisation to impose requirements relating to protection and/or ownership of intellectual property (IP) in research output and how the research output is to be disseminated.

The Australian Government now provides more than \$5 billion annually in funding science and innovation.⁶⁶ In some fields (for example, human health-related biotechnology), virtually all research carried out in Australia (whether in universities, research institutes, or government departments or agencies) is funded by the Australian Government. The Australian government understands that '[a]ccess is a critical issue in the drive to optimise Australia's research infrastructure'.⁶⁷ The Australian Government Department of Education, Science and Training's (DEST) *National Collaborative Research Infrastructure Framework – Strategic Roadmap* (2006)⁶⁸ states:

Consistent with the NCRIS principles, the Roadmap identifies those capabilities that will provide the most strategic impact in terms of delivering national benefit, producing world-class excellence in both discovery and application driven research, and/or enhancing the overall capacity of the research and innovation system by providing enabling research platforms and promoting accessibility and collaboration.⁶⁹

⁶⁶ Australian Government, Productivity Commission, *Terms of Reference for Economic, Social and Environmental Returns on Public Support for Science and Innovation in Australia* (March 2006) <<http://www.pc.gov.au/study/science/tor.html>> at 13 June 2006.

⁶⁷ Department of Education, Science and Training (DEST), *National Collaborative Research Infrastructure Strategy (NCRIS) Strategic Roadmap* (February 2006) 3 <<http://www.dest.gov.au/NR/rduonlyres/E2001074CDA2-4CEA-A1B4-775B4882A5F5/9519/NCRISStrategicRoadmap.pdf>> at 25 July 2006.

For a more recent indication of government policy see: The Hon Kim Carr Minister for Industry Innovation Science and Research, "There is More than One Way to Innovate" 7 Feb 2008 <<http://minister.industry.gov.au/SenatortheHonKimCarr/Pages/>>

⁶⁸ Department of Education, Science and Training (DEST), *National Collaborative Research Infrastructure Strategy (NCRIS)*, 28 February 2006, <http://www.dest.gov.au/sectors/research_sector/policies_issues_reviews/key_issues/ncriis/> at 25 July 2006.

⁶⁹ <<http://www.dest.gov.au/NR/rduonlyres/E2001074-CDA2-4CEA-A1B4775B4882A5F5/9519/NCRISStrategicRoadmap.pdf>> at 25 July 2006.

A critical issue is how to strike the appropriate balance between commercialisation and increased access.⁷⁰ It follows that research funding bodies need to review the terms of their funding agreements to ensure that the objective of providing open access to research results is not contradicted by obligations on funding recipients to protect and commercialise IP that is developed in funded projects.

Promoting Self-Archiving in Open Access Repositories

According to Stevan Harnad:

Articles made 'Open Access', (OA) by self-archiving them on the web are cited twice as much, but only 15% of articles are being spontaneously self-archived. The only institutions approaching 100% self-archiving are those that mandate it. Surveys show that 95% of authors will comply with a self-archiving mandate; the actual experience of institutions with mandates has confirmed this.⁷¹

Since surveys indicate that a majority of researchers favour research funding bodies mandating self-archiving⁷² and, as 95% of authors say

⁷⁰ See generally, Department of Education, Science and Training, *Knowledge Transfer and Australian Universities and Publicly Funded Research Agencies*, March 2006, available at <http://www.dest.gov.au/NR/rdonlyres/36818C20-9918-4729-A150464B662644B3/12630/Knowtran_FinalCompilation_005_web1.pdf> at 30 July 2006.

⁷¹ Stevan Harnad, 'Monitoring Research Impact Through Institutional and National Open-Access Self-Archiving Mandates' in Keith Jeffrey (ed), *Proceedings of CRIS2006. Current Research Information Systems: Open Access Institutional Repositories* (in press) (2006) <<http://eprints.ecs.soton.ac.uk/12093/>> at 16 July 2006; An example is CERN <<http://public.web.cern.ch/Public/Welcome.html>> which has an institutional self-archiving mandate and is close to providing open access to 100% of its own current published research article output in its institutional repository, see <<http://cdsweb.cern.ch/>> at 31 July 2006.

⁷² Sue Sparks, *JISC Disciplinary Differences Report*, Rightscom Ltd (2005) 7 <http://www.jisc.ac.uk/uploaded_documents/Disciplinary%20Differences%20and%20Needs.doc> at 31 July 2006. For a survey of public attitudes to access to publicly funded research output, see Alliance for Taxpayer Access, *Americans Support Free Access to Research*, 31 May 2006, Washington DC <<http://www.taxpayeraccess.org/media/Release06-0531.html>> at 16 July 2006.

they would comply with a self-archiving mandate,⁷³ it has been proposed that institutions and funding bodies should mandate that the author's final draft⁷⁴ must be deposited into the institutional repository immediately upon acceptance for publication.⁷⁵

In recent years, research funding bodies in the United States (US), the United Kingdom (UK) and Germany have adopted open access policies and guidelines calling upon researchers to publish in open access journals and to deposit materials resulting from funded research in an open access repository.⁷⁶

In the US, in February 2005 the National Institutes of Health (NIH), the world's largest non-military research funder, 'prodded by federal departments and Congressional committees', adopted an Open Access Policy⁷⁷ with the aim of increasing the availability of research that it funds. The policy requested all NIH-funded investigators to submit, from 2 May 2005, an electronic version of the author's final, peer-reviewed manuscripts to the PubMed Central⁷⁸ database, the NIH's free

⁷³ Alma Swan and Sheridan Brown, (2005) *Open access self-archiving: An author study*. Technical Report, External Collaborators, Key Perspectives Inc, Table 30, 64 <<http://eprints.ecs.soton.ac.uk/10999/01/jisc2.pdf>> at 13 September 2006.

⁷⁴ That is, the version that is commonly referred to as the 'PostPrint'.

⁷⁵ For the best up-to-date overview of the open access policies applied or being developed by funding bodies in the US, UK, Canada, South Africa and several European countries (including Sweden, France and Germany), focusing on whether open access is mandated or merely encouraged, see Peter Suber, 'Ten Lessons from the Funding Agency Open Access Policies', 2 August 2006, *SPARC Open Access Newsletter*, #100 <<http://www.earlham.edu/%7Epeters/fos/newsletter/08-02-06.htm>> at 3 August 2006; See also Stevan Harnad, 'Monitoring Research Impact Through Institutional and National Open-Access Self Archiving Mandates' in Keith Jeffrey (ed), *Proceedings of CRIS2006. Current Research Information Systems: Open Access Institutional Repositories* (in press) (2006), <<http://eprints.ecs.soton.ac.uk/12093/>> at 16 July 2006; Stevan Harnad, 'Opening Access by Overcoming Zeno's Paralysis' in Neil Jacobs (ed), *Open Access: Key Strategic, Technical and Economic Aspects* (2006) (forthcoming), Self-archived March 19, 2006 <<http://eprints.ecs.soton.ac.uk/12094/>> at 16 July 2006.

⁷⁶ For an overview of research funding bodies' policies on open access, see the European Commission, Directorate-General for Research, 69–70.

⁷⁷ National Institutes of Health, Office of Extramural Research, *Policy on Enhancing Public Access to Archived Publications Resulting from NIH-Funded Research* <<http://publicaccess.nih.gov/>> at 23 May 2006 and <<http://publicaccess.nih.gov/policy.htm>> at 23 May 2006.

⁷⁸ See: <<http://www.pubmedcentral.nih.gov/>> at 23 May 2006.

digital archive of journal literature in the biomedical and life sciences, upon acceptance for publication. The policy applied to any journal articles resulting from research supported wholly or partially with direct funds from NIH. However, in a survey conducted by Janice Hopkins Tanne in 2006,⁷⁹ it was found that less than 5% of NIH-funded researchers were acting in accordance with the NIH's policy.⁸⁰ On 11 January 2008, NIH announced a revision to its Open Access Policy that made its application mandatory rather than voluntary for all peer-reviewed articles arising in whole or in part from direct costs funded by NIH, or from NIH staff, that are accepted for publication on or after 7 April 2008.⁸¹ Funded researchers/institutions were given the responsibility for ensuring that any publishing or copyright agreements concerning submitted articles fully comply with the NIH Open Access Policy.⁸²

The Research Councils UK (RCUK) revised *Position Statement on Access to Research Outputs* of 28 June 2006 (2006 Position Statement)⁸³ endorsed the following principles:

- ideas and knowledge derived from publicly-funded research must be made available and accessible for public examination as rapidly as practicable
- published research outputs should be effectively peer-reviewed

⁷⁹ Janice Hopkins Tanne, 'Researchers funded by NIH are failing to make data available' (2006) 332 *BMJ* 684 (25 March).

⁸⁰ From Peter Suber, *Open Access News* <<http://www.earlham.edu/~peters/fos/fosblog.html>> at 24 March 2006.

⁸¹ See Peter Suber, 'The mandates of January' 118 *SPARC Open Access Newsletter*, 2 February 2008 <http://www.earlham.edu/~peters/fos/newsletter/02-02-08.htm> at 25 March 2008; see also the NIH Public Access webpage: <<http://publicaccess.nih.gov/>>.

⁸² Ibid; see also Kylie Pappalardo, *Understanding Open Access in the Academic Environment: A Guide for Authors*, June 2008, OAK Law Project, available at <http://eprints.qut.edu.au/archive/00013935/01/Microsoft_Word_-_Final_Draft_-_website.pdf>.

⁸³ RCUK, *Position Statement on Access to Research Outputs* <<http://www.rcuk.ac.uk/access/2006statement.pdf>> at 30 July 2006.

- models and mechanisms for publication and access must be an efficient and cost-effective use of public funds
- outputs must be preserved and remain accessible for future generations.⁸⁴

While each of the eight Research Councils (representing diverse research disciplines)⁸⁵ were not directly required to mandate OA archiving for all RCUK-funded research, each were encouraged to develop specific guidelines for the communities it funded, relating to access to research outputs in the particular field/s of research. The intention was to ensure that each discipline was able to respond in ways best suited to its own needs. To date, all but one of the Research Councils have adopted a mandate requiring deposit of peer-reviewed research outputs in an OA repository.⁸⁶ The access policies of the RCUK, along with the policies of research funding bodies in other countries (such as Germany and the US), are included in the Juliet website established by SHERPA.⁸⁷

Similarly in Europe, the European Research Council (ERC) requires that all peer-reviewed publications resulting from funded research be deposited in an openly accessible repository within six months of publication.⁸⁸

⁸⁴ RCUK, News Release, 28 June 2006

<<http://www.rcuk.ac.uk/press/20060628openaccess.asp>> at 30 July 2006.

⁸⁵ There were originally eight Research Councils: Arts & Humanities Research Council (AHRC); Biotechnology & Biological Sciences Research Council (BBSRC); Council for the Central Laboratory of the Research Councils (CCLRC); Economic & Social Research Council (ESRC); Engineering & Physical Sciences Research Council (EPSRC); Medical Research Council (MRC); Natural Environment Research Council (NERC) and Particle Physics & Astronomy Research Council (PPARC). On 1 April 2007, PPARC and CCLRC merged to become the Science and Technology Facilities Council (STFC): see Wikipedia, "Particle Physics and Astronomy Research Council" <http://en.wikipedia.org/wiki/Particle_Physics_and_Astronomy_Research_Council> at 7 July 2008.

⁸⁶ Each of the Research Councils except the Engineering & Physical Sciences Research Council (EPSRC) have adopted OA mandates: see SHERPA-JULIET, Research funders' open access policies <<http://www.sherpa.ac.uk/juliet/>> at 7 July 2008.

⁸⁷ See <<http://www.sherpa.ac.uk/juliet>>.

⁸⁸ Policy accessed via SHERPA-JULIET at: <<http://www.sherpa.ac.uk/juliet>>; see also ERC *Scientific Council Guidelines for Open Access*, 17 December 2007,

In Australia, while 90 per cent of the \$4.3 billion expended on research and development by higher education institutions in 2004 was funded by government,⁸⁹ there is not as yet a policy mandating the archiving/depositing of researching articles in open access repositories. However, Australia's primary funding bodies, the Australian Research Council (ARC) and the National Health and Medical Research Council (NHRMC), moved in 2007 to encourage funded researchers to deposit their research results in open access repositories.⁹⁰ The ARC also requires researchers who are not intending to deposit their research publications in an OA repository to explain their reasons for refraining. This places a greater emphasis on researchers to consider the reasons for their decision and whether those reasons are justifiable.

Government reports by the Australian Law Reform Commission (ALRC): *Genes and Ingenuity: Gene Patenting and Human Health*,⁹¹ and DEST: *Review of Closer Collaboration between Universities and Major Publicly Funded Research Agencies* (2005)⁹² and *Analysis of the Legal Framework for Patent Ownership in Publicly Funded Research Institutions* (2003)⁹³ while not focusing directly on the question of imposition of conditions regarding

<http://erc.europa.eu/pdf/ScC_Guidelines_Open_Access_revised_Dec07_FINAL.pdf> at 25 March 2008.

⁸⁹ Buchhorn and McNamara, *Issues for Research Data*, 26.

⁹⁰ Australian Research Council (ARC), *Discovery Projects Funding Rules for funding commencing in 2009*, 13 <http://www.arc.gov.au/ncgp/dp/dp_fundingrules.htm> at 25 March 2008; National Health and Medical Research Council, *NHMRC Project Grants Funding Policy for funding commencing in 2009*, 23

<<http://www.nhmrc.gov.au/FUNDING/apply/granttype/projects/index.htm>> at 25 March 2008.

⁹¹ Australian Law Reform Commission, *Genes and Ingenuity: Gene Patenting and Human Health*, ALRC Report No 99 (2004)

<<http://www.austlii.edu.au/au/other/alrc/publications/reports/99/>> at 30 May 2006.

⁹² Department of Education, Science and Training, *Review of Closer Collaboration between Universities and Major Publicly Funded Research Agencies* (2004)

<<http://www.dest.gov.au/NR/rdonlyres/327F4C1D-99CC4F93-91FB-1A2DEA8F299E/3623/pub.pdf>>.

⁹³ Department of Education, Science and Training, *Analysis of the Legal Framework for Patent Ownership in Publicly Funded Research Institutions* (2003)

<http://www.dest.gov.au/sectors/research_sector/publications_resources/other_publications/patent_ownership_in_publicly_funded_research_institutions.htm#6._Recommendations_for_Australia> at 23 May 2006.

OA to publications and other output resulting from funded research projects are significant in that they demonstrate that the issue of attaching conditions to funding grants to ensure that the project output is dealt with in the desired manner and, in particular, is consistent with the funding body's public benefit objectives.

A 2006 report to DEST titled *Research Communication Costs in Australia - Emerging Opportunities and Benefits* recognises the importance of conditioning grants in promoting access to research results:

Research evaluation is the primary point of leverage, influencing strongly the scholarly communication and dissemination choices of researchers and their institutions. A related secondary point of leverage is funding, and the conditions funding bodies put upon it. To attain the goals of accessibility articulated in the Accessibility Framework (Appendix III) and elsewhere, and realise the potential benefits of enhanced access, it will be essential to ensure that funding and grant assessment, research evaluation and reward take account of emerging possibilities and opportunities, and build in open access options.⁹⁴

The Accessibility Framework referred to is the Quality and Accessibility Frameworks for Publicly Funded Research⁹⁵ proposed by the Australian Government in May 2004 as part of the *Backing Australia's Ability - Building our Future through Science and Innovation*.⁹⁶ The Accessibility Framework is intended to provide a strategic framework to improve access to research information, outputs and infrastructure.⁹⁷

As part of the policy development process, universities and research funders need to closely consider the benefits of open access to knowledge. The Open Access to Knowledge (OAK) Law Project,

⁹⁴ John Houghton, Colin Steele and Peter Sheehan, *Research Communication Costs in Australia. Emerging Opportunities and Benefits*, (September 2006) 7
<http://www.dest.gov.au/NR/rdonlyres/0ACB271F-EA7D4FAF-B3F7-0381F441B175/13935/DEST_Research_Communications_Cost_Report_Sept2006.pdf> at 10 October 2006 (hereinafter Houghton, Steele and Sheehan, *Research Communication Costs*).

⁹⁵ See <http://backingaus.innovation.gov.au/2004/research/qual_pub_res.htm>.

⁹⁶ See <<http://backingaus.innovation.gov.au/>>.

⁹⁷ Houghton, Steele and Sheehan, *Research Communication Costs*, 132.

funded by DEST under the Systemic Infrastructure Initiative, is seeking to provide institutions and research funders with assistance in identifying these benefits and guidance to promote the adoption of effective and cutting-edge copyright management frameworks.⁹⁸

B. Author – Employer (Employment Agreement and IP Policy)

Universities and research institutes may require their academic and research staff to make their academic and research output available through OA institutional⁹⁹ or disciplinary (or subject-based)¹⁰⁰ repositories. The legal context in which this outcome is secured is the relationship between the university or research institute as employer and the academic or research project author as employee.¹⁰¹

Since the mid-1990s, the majority of Australian universities have developed IP policies addressing ownership of IP (patents, copyright, confidential information etc.) generated in the course of academic or research activities performed within the scope of the employment relationship. Intellectual Property policies are often part of the formal regulations approved by the governing body of the university for its administration and are generally published in the university handbook and on the institutional website. Such policies may also be incorporated

⁹⁸ See Professor Brian Fitzgerald et al, *The OAK Law Project Report No. 1 - Creating a legal framework for copyright management of open access within the Australian academic and research sector* <<http://www.oaklaw.qut.edu.au>> at 28 September 2006.

⁹⁹ See Harvard University Faculty of Arts and Sciences, *Faculty of Arts and Sciences Agenda* (February 2008) <http://www.fas.harvard.edu/~secfas/February_2008_Agenda.pdf> at 25 March 2008. Institutional repositories assist in raising the profile of institutions, making their research output more visible and accessible.

¹⁰⁰ Disciplinary, or subject-based, archives provide efficient and centralised access to full text articles in specific domains. Eight disciplines have successfully set up e-print archives: high-energy physics and mathematics (arXiv), economics (RePEc), cognitive science (CogPrints), astronomy, astrophysics and geophysics (NTRS and ADS), and computer science (NCSTRL).

¹⁰¹ For case law addressing the issue of whether an institution can enforce university policies (in the context of patent ownership) through a faculty member's employment contract (by reference, either specific or general, to the policies in the employment contract) see: *Victoria University of Technology v Wilson and Others* [2004] VSC 33 and *University of Western Australia v Gray* (No 20) [2008] FCA 498.

by reference into employment contracts between the university and its employees.

A range of approaches to the question of copyright ownership can be found in university IP policies. Most policies seek to balance the interests of the parties by reserving certain rights to the party which does not own copyright. In a review of university copyright policies, the Zwolle project identified the following three approaches taken by UK universities:¹⁰²

Scenario A: individuals own copyright with a licence to the institution
University College London, UK: ‘UCL recognises the rights of its staff to ownership of copyright in research publications, books and other similar academic publications in all format ... UCL will seek to secure, free, unconditional and perpetual, non-exclusive licence to use academic and teaching materials in all formats which are generated by members of staff arising out of employment by UCL.’ The policy is available at www.ucl.ac.uk/staff/resources/copyright-policy/

Scenario B: institution owns copyright but university agrees not to benefit from individuals’ work

University of Bristol, UK: ‘University policy adopts and imposes UK Statute [Copyright, Designs and Patents Act 1988]. University policy is set out in the Standing Orders of Council e.g. section 12.3 of the Standing Orders of Council governing the appointment of members of the Non-professorial Academic Staff. Normally, therefore, the University is the first owner of IP and IP rights generated by its employees ... The University will not in normal circumstances seek to benefit from any rights it may have as employer in the academic publications of members of the Academic Staff.’ The policy is available at

<<http://www.bris.ac.uk/research/ip/policy/ownership.html>>

¹⁰² Zwolle Group, ‘Implementing the Zwolle Principles: University Copyright Policies’
<http://www.surf.nl/copyright/keyissues/scholarlycommunication/implementing_policies.php>

at 16 July 2006 (hereinafter Zwolle Principles).

Scenario C: institution owns IP rights but publications excepted or rights waived

University of Oxford, UK: 'The University claims ownership of all IP ... devised, made or created ... by persons employed by the University in the course of their employment ... Notwithstanding section 6 of this statute, the University will not assert any claim to the ownership of copyright in ... artistic works, books, articles, plays, lyrics, scores, or lectures, apart from those specifically commissioned by the university.' The policy is available at

<<http://www.admin.ox.ac.uk/rso/policy/ip.shtml>>

The IP policies adopted by Australian universities typically vest ownership of copyright in some materials (e.g. course guides and handbooks) in the university while providing for copyright in a wide range of other materials (including published journal articles, books and reports) to be owned or controlled by the employee author/s.¹⁰³ This splitting of copyright according to the nature and purpose of the material is apparent in many university IP policies.

An example is Charles Sturt University's IP policy¹⁰⁴ which states that the university owns all IP created by an employee author in pursuance of the author's duties under a contract of employment, including copyright in 'courseware (books, print, videos, CD-ROMs, manuals, audiovisual recordings, computer software or other materials) created specifically for use in, or in connection with, a course, subject or unit offered by the university'.¹⁰⁵ On the other hand, employee authors own IP in copyright works 'the subject matter of which is primarily concerned with scholarship, research, artistic expression, creativity or academic debate', including 'books, articles or other similar works, whether in written or any other form', 'artistic works created by researchers in fine art or design' and 'any other professional work' created by an employee

¹⁰³ For a comprehensive overview of the IP policies of Australian universities, see Monotti and Ricketson, *Universities and Intellectual Property*. The IP policies of many Australian universities are set out on the SURF website

<<http://www.surf.nl/copyright/keyissues/countries/australia.php>> at 16 May 2006.

¹⁰⁴ Version 4.0, adopted 1997, last modified August 2007,

<<http://www.csu.edu.au/adminman/tec/PolicyonIntellectualProperty.pdf>> at 16 May 2006.

¹⁰⁵ Clause 6.1.

author.¹⁰⁶ The policy expressly excludes from employee copyright ownership materials which ‘were prepared for CSU course work and teaching’, ‘were created using IP owned by CSU’ or if CSU ‘has made a specific and significant contribution of funding, resources, facilities or apparatus which led to the creation of [the] works’.¹⁰⁷

For those materials in which the IP policy vests copyright ownership in the university, no problems arise. The university, as copyright owner, can exercise all the rights required to make the material available through its own institutional open access repository or an external disciplinary repository. By contrast, where the terms of the IP policy vest copyright ownership in the employee, the situation is more complex and needs to be carefully managed by the university if it is to ensure that its employees do not, by exercising their rights as copyright owners, limit the university’s ability to implement its policy on open access to academic and research output. In particular, in the absence of any restriction imposed by the university (whether through its IP policy or express terms of the employment contract) there is nothing to prevent employed academics and researchers who own copyright in their academic and research output from assigning copyright or granting an exclusive licence to a third party (such as a publisher), without reference to the university. In the typical case, where the assignment is of the whole of the copyright (for example, in the traditional Publishing Agreement), the university will not be in a position to require the material to be made available in an institutional or disciplinary repository once the transfer has been effected.

Where a university is seeking to develop a comprehensive OA institutional repository containing the academic and research output of its employees, it should review the terms of its employment contracts and IP policy to ensure consistency between the institution’s policies regarding OA to academic and research output and the obligations imposed on academic and research staff. To address the problems arising from copyright transfer by employees, it may be appropriate for universities to include in their IP policies a requirement that before transferring copyright ownership to a third party the employee must first grant the university all the rights required to enable it to make the

¹⁰⁶ Clause 6.2, paras (a) to (d).

¹⁰⁷ Clause 6.2, paras (e) to (h).

material available in an OA repository. Such grant of rights may take the form of an assignment of part of the copyright to the university or it may be in the form of an irrevocable, non-exclusive licence in favour of the university. In either case, it should expressly state the rights granted to the university and should be in writing, signed by the employee.

The first Australian university to implement a formal requirement for academic authors to deposit all academic and research output was the Queensland University of Technology (QUT), under *Policy F/1.3 E-print repository for research output at QUT*,¹⁰⁸ adopted in 2003.¹⁰⁹ The QUT E-prints policy states that deposit of materials is subject to ‘any necessary agreement with the publisher’ and advises that ‘guidance on copyright arrangements and standards for publishers is available from the University Copyright Officer’. The deposit policies of all other Australian universities (except CSU and UTas) are based on voluntary submission by academic and research staff. In a recent survey conducted by Professor Arthur Sale of the proportion of DEST¹¹⁰ funded research output deposited in institutional repositories, it was found that no Australian university with a voluntary policy collects significantly more than 15% of DEST reportable content, and in most cases the amount was considerably less. This finding was comparable with international surveys which have also found 15% to be the average deposit level achieved voluntarily.¹¹¹ In comparison, QUT achieved deposit rates 2.5 times higher than its nearest competitor in 2004 and 5 times higher in 2005, with estimated deposit rates of 60% for 2005 and 80% for 2006.¹¹² Sale attributes the difference between the high deposit levels being achieved by QUT as compared to those observed at other Australian universities to ‘the deposit policy coupled with good author support

¹⁰⁸ <http://www.mopp.qut.edu.au/F/F_01_03.html> at 8 May 2006.

¹⁰⁹ Since that time, Charles Sturt University has implemented a mandatory deposit policy for all staff (in January 2008), and the University of Tasmania has been implementing a university-wide deposit mandate in a “patchwork” fashion – department by department. The School of Computing at University of Tasmania has had a deposit mandate in place since 2006.

¹¹⁰ Commonwealth Department of Education, Science and Technology.

¹¹¹ A. Sale, “Comparison of Content Policies for Institutional Repositories in Australia” (2005) <http://www.firstmonday.org/issues/issue11_4/sale/> at 27 July 2008

¹¹² A. Sale, “Comparison of Content Policies for Institutional Repositories in Australia” (2005), 3 <http://www.firstmonday.org/issues/issue11_4/sale/> at 27 July 2008

practices’,¹¹³ a finding consistent with a major international study by Swan and Brown in 2005.¹¹⁴ Sale drew the following conclusion:

A requirement to deposit research output into a repository coupled with effective author support policies works in Australia and results in high deposit rates ... Authors are willing to comply with a requirement to deposit. Voluntary deposit policies do not result in significant content, regardless of any author support ...¹¹⁵

It should also be noted that recent developments, most notably at Harvard University, have seen university academics vote to subject themselves to a requirement to provide their university with permission to make their scholarly articles available in an institutional open access repository. The Harvard Faculty of Arts and Sciences adopted such a policy on 12 February 2008.¹¹⁶

C. Author – Publisher (Publishing Agreement)

The degree of control that an academic author is able to exercise in respect of a published article, in terms of the use that the author can personally make of it or authorise others to make of it, depends on the scope of the rights (if any) that the author has in the published article. This, in turn is largely dictated by the legal relationship between the author and publisher, as established by the Publishing Agreement. The extent to which authors of published articles can continue to reproduce, distribute or provide access to the article, for example, by self-archiving

¹¹³ Sale, *Comparison of Content Policies*, 3.

¹¹⁴ Alma Swan and Sheridan Brown, *Open access self-archiving: an author study*, Technical Report, External Collaborators, JISC, HEFCE <<http://eprints.ecs.soton.ac.uk/10999>> at 16 July 2006.

¹¹⁵ Sale, *Comparison of Content Policies*.

¹¹⁶ For the Harvard Faculty of Arts and Sciences policy, see: <http://www.fas.harvard.edu/~secfas/February_2008_Agenda.pdf> at 25 March 2008; for more information about the Harvard Faculty of Arts and Sciences policy, see Peter Suber, “The open access mandate at Harvard” *SPARC Open Access Newsletter*, Issue #199, 2 March 2008, <<http://www.earlham.edu/~peters/fos/newsletter/03-02-08.htm>> at 20 May 2008; see also, Kylie Pappalardo, *Understanding Open Access in the Academic Environment: A Guide for Authors*, June 2008, OAK Law Project, available at <http://eprints.qut.edu.au/archive/00013935/01/Microsoft_Word_-_Final_Draft_-_website.pdf>.

it or depositing it with an institutional or disciplinary repository, depends on the scope of the rights (if any) retained by the author.

Even though the author has written the article, if they have assigned copyright to the publisher and have not obtained a licence back from the publisher permitting them to continue reproducing and distributing the article, their actions in doing so will be every bit as much an infringement of copyright as if the acts were done by a completely unrelated third party. Likewise, if academic writers are to permit third parties to use their published articles, they must have the authority to be able to grant that permission. In particular, academic authors who wish to submit copies of their published articles to digital repositories from which they can be reproduced, viewed etc. by the public at large or by members of a qualified community, must be able to warrant to the repository manager ("custodian") that they have the rights to authorise the repository to make the copyright material available to those who access the repository. That is, they have the rights to reproduce, first publish and communicate electronically the copyright material to the public (in other words, by making the material available on a website or by transmitting the material in digital form).

Much of the discussion of the allocation of rights between publishers and authors in the academic context has started from the assumption that copyright is assigned in its entirety from the author to the publisher at the time the publishing arrangements are agreed. There has also been little discussion of the importance of identifying the actual owner of copyright in a published article. Too often, discussion of authors' rights in relation to ongoing use of their published articles has been based upon assumptions which do not necessarily apply across the board. There has been a tendency to assume that the author has, prior to publication, assigned copyright to the publisher. The focus on the publisher as controlling the ongoing use of published articles has tended to put into the shadows alternative models of rights management, which involve a lesser ceding of control by the author, (for example, through a partial transfer of copyright or merely granting the publisher a licence to publish). If the participants in the discussion were to shift their focus they would find that the increased emphasis on OA has been accompanied by a shift away from the dominant model in favour of one in which copyright is retained by the author, the publisher is granted a

licence to publish and the author retains rights over further re-use of the material.

The range of models of copyright management in the author-publisher relationship, can be seen along a continuum of control, with maximum control by the author at one extremity and maximum control by the publisher at the other. At the one end of the spectrum the author retains copyright (and thereby maximum control) and merely licenses the publisher to publish the article, on an exclusive, sole or non-exclusive basis. At the other end of the spectrum, the publisher obtains a full assignment of copyright from the author (and thereby maximum control) and does not permit the author to self-archive the article (either in its draft PrePrint form or the published PostPrint) or further distribute it (although the author may purchase hard copy reprints). In retaining copyright the author has control of further distribution of the article (including the right to self-publish, self-archive or deposit it in a repository).

Points along the continuum from maximum author control to maximum publisher control can be identified, in broad terms, as follows:

1. Author retains copyright and controls distribution (which may include self-publishing, self-archiving or depositing it in a repository).
2. Author retains copyright and grants a licence (exclusive, sole or non-exclusive) to publisher to publish the article.

Ultimately in developing any licensing model for managing the author-publisher relationship, the scope of the rights granted to the publisher will be determined by how the licence deals with a range of issues, including:

- whether the licence granted is exclusive, sole, or non-exclusive
- the period of time for which the licence is granted
- the territory covered by the licence
- whether any restrictions are imposed on the commercial use of the material (or whether it can be used only for non-commercial purposes)
- conditions applying to any further distribution of the material.

3. Author assigns copyright partially to publisher, retaining (reserving) ownership of part of the copyright.
4. Author assigns copyright to publisher but obtains an express licence back from publisher to further reproduce and distribute, on terms determined by publisher.
5. Author assigns copyright entirely to publisher, with an implied licence to self-archive or deposit the article into an institutional or disciplinary repository.
6. Author assigns copyright entirely to publisher.

Recent surveys of authors have clearly indicated a preference for a copyright model under which the author retains copyright and continues to be able to exercise rights over re-use of the material for educational, academic or commercial purposes.¹¹⁷ In *The Institutional Repository* (2006), Jones, Andrew and MacColl comment that they have ‘noted that the major difficulties with clearing permission arise when dealing with materials that are not owned by the submitting author [and] advocate that [generally speaking] authors should retain as much of their rights as possible’.¹¹⁸

Examples of Rights Management Models under the Author-Publisher Relationship

1. Author Retains Copyright and Grants a Licence (Exclusive, Sole or Non-Exclusive) to Publisher to Publish the Article.

This is the model favoured by the Open Access Law Program established by Lawrence Lessig (Stanford Law School), Michael Carroll (Villanova Law School) and Dan Hunter under the umbrella of the Science Commons project.¹¹⁹ It encourages authors to negotiate

¹¹⁷ See Maurits van der Graaf and Esther Hoorn, *Towards good practices of copyright in Open Access Journals* (2005) the first output of the JISC-SURF partnering on copyright project <<http://www.surf.nl/en/publicaties/index2.php?oid=50>> at 26 May 2006; see also, Anthony Austin, Maree Heffernan and Nikki David, *Academic Authorship, Publishing Agreements & Open Access: Survey Results*, March 2008, OAK Law Project, available on the OAK Law website at <<http://www.oaklaw.qut.edu.au>>.

¹¹⁸ Richard Jones, Theo Andrew and John MacColl, *The Institutional Repository* (2006) 54–155.

¹¹⁹ See <<http://sciencecommons.org/literature/oalawpublication>>.

individually with the journals in which they publish, to retain ownership of copyright and the right to deposit their material in OA repositories. The program has developed the following resources to promote OA in legal publishing, including:

- *The Open Access Law Journal Principles*: The OAL Program encourages law journals to commit to a set of OAL Journal Principles. These Principles require that a journal: (1) take only a limited term licence, (2) provide a citable copy of the final version of the article, and (3) provide public access to the journal's standard publishing contract. In return, the author promises to attribute first publication to the journal. (See <http://sciencecommons.org/literature/oalawjournal>).
- *The Open Access Law Author Pledge*: For authors wishing to commit publicly to OA ideals, we have established an OAL Author Pledge. This pledge commits authors to only publish law review articles in journals that adhere to a minimum OAL commitment.
- *The Open Access Model Publishing Agreement*: The OAL Program also provides a Model Agreement that embodies the OAL Journal Principles in a fair and neutral contract that is easy for both authors and law reviews to adopt. It also provides for an easy mechanism for authors and journals to adopt Creative Commons (CC) licences to make their work more easily available.¹²⁰

2. Author Assigns Copyright Partially to Publisher, Retaining (Reserving) Ownership of Part of the Copyright.

Under this model, based on the splitting of copyright interests among the parties, the author assigns copyright partially to the publisher but retains (or 'reserves') certain key rights required to enable them to control certain uses of the article, for example, to enable the author to self-archive the article or to deposit it in a digital repository.

¹²⁰ See <<http://sciencecommons.org/literature/oalawpublication>>.

This model underlies the so-called SPARC Author Addendum (or simply, SPARC Addendum)¹²¹ developed by Professor Michael Carroll of Villanova Law School for the Scholarly Publishing and Academic Resources Coalition (SPARC).¹²² The SPARC Addendum is a set of clauses intended for inclusion by an author in a standard Publication Agreement in which copyright is assigned to the publisher, in order to limit what would otherwise be a general transfer of copyright, by excluding from the transfer certain distribution rights which are reserved to the author. In particular, the SPARC Addendum reserves to authors certain key rights, in particular, the right to post their articles in digital repositories.¹²³

3. Author Assigns Copyright to Publisher but Obtains an Express Licence Back from Publisher to Further Reproduce and Distribute, on Terms Determined by Publisher.

The prevalence of the copyright assignment model is apparent from the survey of publishers conducted by the UK SHERPA (Securing a Hybrid Environment for Research Preservation and Access) project. The information about publishers' practices on the SHERPA website¹²⁴ shows that the majority obtained a transfer of copyright from the author. The SHERPA website provides a useful overview of publishers' practices, with a primary focus on whether or not they permit authors to self-archive or further distribute pre-prints and post-prints.

In formulating the SHERPA categorisation (green/blue/yellow/white) much emphasis was placed on the policies issued by publishers. Such policies represent to the public at large the publisher's practices. In some cases, for example, where the publisher's policy states that authors are permitted to self-archive, or make the published article available in an institutional or disciplinary repository, the publisher may be going

¹²¹ Version s 2.1 of the SPARC Addendum is available at <http://www.arl.org/sparc/author/docs/AuthorsAddendum2_1.pdf> at 2 August 2006.

¹²² See: <<http://www.arl.org/sparc/>> at 16 July 2006.

¹²³ See: <<http://www.arl.org/sparc/author/addendum.html>> and <http://www.arl.org/sparc/author/docs/AuthorsAddendum2_1.pdf> at 16 July 2006. For a guide to the SPARC Author Addendum, see SPARC, *Author Rights* <<http://www.arl.org/sparc/author/addendum.html>> at 2 August 2006.

¹²⁴ <<http://www.sherpa.ac.uk/>> at 2 July 2006.

beyond what has been expressly stated in their standard, written publishing agreements which provide for assignment of copyright by the author but are silent as to any rights the author may have to further use or distribute the published article. In this case, the question arises as to whether the general statement of policy can be regarded as unilaterally varying the express terms of the existing publishing agreements with authors. The more likely situation is that the publishers' policy statements are merely a representation which, if acted on by authors, cannot be disavowed by publishers (doctrine of estoppel). Essentially, the publisher is indicating that it will not enforce its rights as copyright owner, if the author makes use of the published article in the manner described by the publisher in its policy statement.

While the publishers' policy statements have retrospective effect in relation to existing contracts, it would be expected that new contracts would be drafted to expressly reflect the published policy.

4. Author Assigns Copyright Entirely to Publisher, With an Implied Licence to Self-Archive or Deposit the Article into an Institutional or Disciplinary Repository.

Many publishers require the author to assign copyright and, while the question of the author's rights to self-archive or deposit the article (in pre-or post-print version) is not expressly addressed in the Publishing Agreement, the circumstances may give rise to an implied licence to the author to use the article in this way. While there may be circumstances which can be relied upon to support the existence of an implied licence, there will inevitably be uncertainty about the terms and extent of any such licence.

5. Author Assigns Copyright Entirely to Publisher.

Under the traditional model of academic publishing, the author assigns the whole copyright to the publisher in exchange for having the article or work published. Few, if any rights are licensed back to the author.

In the context of pursuit of OA objectives, this option is the least suitable. It minimises the author's control over the published article, while maximising the publisher's ability to prohibit or impose restrictions on further distribution and educational uses of the published work, without consulting the author.

D. Author – Digital Repository (Repository Deposit Licence)¹²⁵

The relationship between the author (or another party who owns copyright in the work, such as the author's employer or the publisher to which copyright has been assigned) and the digital repository in which a copy of the article is deposited is governed by the terms of the Repository Deposit Licence between the parties.

The Repository Deposit Licence will be entered into by the administrator of the digital repository and the author, the author's employer or the publisher.

If the repository is an institutional repository or disciplinary repository established by the author's employing institution, the parties to the Repository Deposit Licence will be the author and their employer.

Surprisingly, many ePrint repositories do not enter into formal agreements with authors who deposit their works because such agreements are thought to discourage authors from depositing. In a 2000 survey of e-print repository practices, the Rights METadata for Open Archiving (RoMEO) project found that about 32% of respondents took it on trust that the author had the right to deposit the work without explicitly asking them to confirm that they held all necessary rights.¹²⁶ However, a 2005 report commissioned by SHERPA on deposit licences for e-prints emphasised the value of such licences in establishing a formal relationship between the repository and authors depositing their works into digital repositories. It concluded that:

[d]eposit agreements should be considered an essential part of an e-print repository's operation. ... For the repository, it provides a formal framework that defines what the repository can and cannot do, making it easier to manage the e-print in the long-term while helping to reduce its legal liabilities. For the author, it provides reassurance that the repository is not

¹²⁵ See further, Kylie Pappalardo and Dr Anne Fitzgerald, *A Guide to Developing Open Access Through Your Digital Repository*, September 2007, OAK Law Project, including Sample Repository Deposit Licence, available at <<http://eprints.qut.edu.au/archive/00009671/01/9671.pdf>>.

¹²⁶ The RoMEO study is referred to in Gareth Knight (Arts & Humanities Data Service), *Report on a deposit licence for E-prints*, 7 (21 June 2004) 1 <http://www.sherpa.ac.uk/documents/D4-2_Report_on_a_deposit_licence_for_E-prints.pdf> at 16 May 2006 (hereinafter Knight, *Deposit Licence Report*).

taking ownership of their work, and makes them aware of what type of service the repository is providing.¹²⁷

It is necessary for a digital repository to determine the basis on which repository content may be accessed and re-used by end users. The Repository Deposit Licence between the author (or publisher) and the repository should address the extent to which the deposited material can be made available to other users and institutions and should grant an express licence to the repository to enable the repository to do all acts required to make the material available for access, use and/or further distribution by end users.

In particular, the matters addressed in the Repository Deposit Licence may include:

- *permissions granted by author (or other copyright owner) to digital repository*
 - grant of a non-exclusive licence to the digital repository
 - extent of rights granted to digital repository, for example, to reproduce, distribute the deposited material (including the abstract) worldwide in print and electronic format in any medium
 - retention by author of rights to make use of the current and future (revised) versions of the deposited work
 - rights granted to digital repository to translate the deposited work (without changing the content) to any medium or format for the purpose of preservation
 - requirement for citation to published version to be included and to be clearly visible
 - author's rights to provide updated versions of the work
 - conditions under which the repository administrators can remove the deposited work.

¹²⁷ Knight, *Deposit Licence Report*.

- rights granted to digital repository to copy the deposited work for purposes of security, back-up and preservation.
- *access to work by other parties*
 - basis on which work is to be made available to other users and institutions
 - rights of other parties to access, use and further distribute the work.
- *representations and warranties by the author (or copyright owner) to repository administrators*
 - representation by author of authority to enter into the Repository Deposit Licence
 - representation by author of right to grant the rights to the digital repository as stated in the Repository Deposit Licence
 - where the deposited work has been sponsored or supported by another organisation, a representation by the author that obligations required by the agreement with such sponsor regarding use of the work have been fulfilled
 - warranty by author that the work is original, and to the best of his or her knowledge, does not infringe any other party's copyright
 - representation that, where the deposited work contains material for which the author does not hold copyright, the author has obtained the unrestricted permission of the copyright owner to grant the digital repository administrator the rights required by the Repository Deposit Licence and that any third party owned material is clearly identified and acknowledged within text or content of the deposited work clearly identified and acknowledged within text or content of the deposited work.

- *responsibility for enforcement of IP*
 - whether administrators of digital repository have any obligations to take legal action on behalf of the author (or copyright holder) in the event of breach of IP rights in the deposited work.

*E. Digital Repository – End Users*¹²⁸

The Repository Distribution (End User) Agreement grants rights to end users to access and re-use the deposited material that are consistent with (and do not extend beyond) the licence granted to the repository by the author (or publisher) under the Repository Deposit Licence.

End users may be individual members of the public or members of a specific academic community with defined access rights. The terms and conditions governing access to and use of material in the repository should be clearly displayed on the repository web site and brought to the attention of end users so they understand that their use of the repository and materials in it is subject to those terms and conditions. In particular, any limits on the rights of end users to copy and further distribute the material in the repository should be stated.

Where it is essential to obtain assent by end users to comply with restrictions on access and use, the click-wrap format should be used for the Repository Distribution (End User) Agreement.¹²⁹ A click-wrap web site agreement involves end users first viewing the terms and conditions governing access to and use of the materials in the repository, and clicking an ‘I accept’ or ‘I agree’ button or icon to indicate that they assent to those conditions before they are able to obtain access to and use articles in the repository. Where restrictions apply and the

¹²⁸ See further, Kylie Pappalardo and Dr Anne Fitzgerald, *A Guide to Developing Open Access Through Your Digital Repository*, September 2007, OAK Law Project, including Sample Repository Deposit Licence, available at <<http://eprints.qut.edu.au/archive/00009671/01/9671.pdf>>.

¹²⁹ A similar approach to that described in this paragraph is advocated by Richard Jones, Theo Andrew and John MacColl in *The Institutional Repository* (2006) 152–4.

repository will not permit access unless end users have agreed to be bound by the terms and conditions of access and use, end users who do not accept the terms and conditions should be given the opportunity of declining (by clicking a 'I decline' or 'I do not agree/accept' button), in which case they will not be permitted to continue to access the repository or download material from it.

In cases where few, if any, restrictions are imposed on access to and use of the materials in the repository, it will suffice if the Repository Distribution (End User) Agreement is in browse-wrap form or if the terms and conditions are available by clicking on hypertext links at the bottom of the repository web site pages. In the browse-wrap form of agreement, the end user is required to view the terms and conditions but is not required to click on a button to indicate assent.

F. Author/Publisher – End Users

Where the article is distributed by the author or publisher (or another copyright holder), the rights of end users are governed by the terms of the Distribution Agreement. If the author has assigned copyright to a publisher, the rights of end users will be determined by the terms of the licence granted to end users by the publisher. However, in cases where the author has retained copyright wholly or partially, it may be the author who directly authorises end users to use the article (Author Distribution Agreement).

An example of an author-end user agreement is the SCRIPT-ed Open Licence ('SOL')¹³⁰ used by the SCRIPT-ed online law journal which takes the form of a non-exclusive licence granted by the author to 'Users'.¹³¹ Users are given the right to disseminate the original and unmodified work, provided it is not done for commercial purposes.¹³²

¹³⁰ See SCRIPT-ed Open Licence (SOL).

¹³¹ *User* is defined as 'the person who reads, copies, issues copies of the work, translates, displays, performs or broadcasts the Work'.

¹³² See <<http://www.law.ed.ac.uk/ahrb/script-ed/policies.asp>> at 16 July 2006. Clause 4 deals with Modification and Clause 5 deals with Adaptations.

G. Copyright Collecting Society – Digital Repository and End Users

In establishing a system to enable access to academic and research materials in online repositories, it is necessary to consider how such materials will be treated under the statutory licence for reproduction and communications of works in electronic form under Division 2A of Part VB of the *Copyright Act*.

The question is whether the obligation to pay remuneration to a collecting society for the use of the copyright work still remains when a licence to use the work is granted expressly or impliedly by the copyright owner. If the obligation to pay remuneration continues in force unless expressly excluded by the terms of the licence to archive the material, this will have implications for the drafting of Publication Agreements.

COPYRIGHT MANAGEMENT ISSUES FOR ELECTRONIC THESES AND DISSERTATIONS

The electronic distribution of theses also raises many copyright issues.

Ownership Principles – the Legal Status of Theses

Copyright

Theses and dissertations will automatically be protected by copyright as a literary work, with the rights vesting in the author who has created them. It should also be noted that a thesis may consist of more than simply literary works or dramatic, musical or artistic works.¹³³ For example, sound recordings and cinematograph films are now common in theses in some disciplines and these materials may also contain more than one layer of copyright. For example, the underlying rights in the script or any sound recording may co-exist alongside the copyright in the film.

Ownership of Copyright in Theses

Subject to any express agreement to the contrary (such as an agreement assigning copyright to the university or a third party), PhD students will own copyright in the original expressions in their theses.

¹³³ *Copyright Act* s 31.

Where a student is receiving a scholarship or there has been a significant investment made towards the student's thesis, the investor may seek to obtain ownership of copyright in the thesis.¹³⁴

Performers' and Moral Rights

Performers' rights may be relevant for theses and dissertations, in particular theses in the area of creative industries and performing arts. In addition to having the personal right to prevent the making, copying or public performance of an unauthorised recording or communication of a live performance (as outlined in 'Overview of Principles of Copyright Law', above), performers also have new economic rights to the extent that the performer who performed the performance and the person at the time of the recording who owned the record (being the person who owned the master recording on which the record was made) and are now co-owners of the copyright in equal shares in the sound recording of the live performance.¹³⁵ These rights are relatively new, following to amendments to the *Copyright Act* arising out of the Australia-United States Free Trade Agreement (AUSFTA).¹³⁶ Performers may assign their share of the copyright to the original copyright owner in the sound recording or to a third party. The normal employment provisions under the *Copyright Act* will also apply — for example, copyright in a performance done in the course of employment will be owned by the employer.

PhD students and researchers could also have moral rights in their theses, including the right to be attributed/cited as the author of a work in third party papers and publications reproducing parts of their thesis. In addition, the moral right of integrity may be relevant for theses in the creative industries, such as film-making or sound production, where the remixing and re-use of aspects of a work (such as in a pastiche or multimedia work) could potentially subject the work to derogatory treatment in a way that demeans the creator's reputation if done without the consent of the creator — thereby infringing their moral right of integrity.

¹³⁴ Ann Monotti and Sam Ricketson, *Universities and Intellectual Property* (2003) Chapter 7 (hereinafter 'Monotti and Ricketson, *Universities and Intellectual Property*').

¹³⁵ *Copyright Act* s 97 (2A). See also ss 100AA-100AH.

¹³⁶ *Copyright Act* s 22 (3A). See ss 22 (3B)-(3C).

A History of the Distribution of Theses

The Pre-Digitisation of Theses

Prior to the digitisation of theses, the thesis service that libraries could provide was necessarily limited. Theses were predominantly distributed in hard copy form, usually a bound copy, which would then be deposited in the library of the degree awarding institution, and perhaps that of the external assessor's institution. The core problem prior to digitisation of theses was that, in the majority of cases theses were not published on a commercial basis. This made it extremely difficult to locate and access theses in many cases, as they were held at the library of the institution where the degree was awarded, with access limited to personal inspection of the hard copy within the library.

In some cases, copies of theses and dissertations are also held in the various state libraries and the National Library of Australia (NLA). However, as the NLA currently does not receive a copy of every thesis awarded by an Australian university it recommends that the relevant institution where the thesis was completed be consulted in order to obtain access to the required theses or dissertations.¹³⁷

In contrast to Australia, the British Library provides a thesis service, which is known as the British Thesis Service, comprising:¹³⁸

- Full text access to over 170 000 doctoral theses dating from the 1970s to today, with most UK universities making their students' theses available on the service.
- This collection of theses is held in either paper bound copies or on microfilm. The service also makes available for sale the majority of theses in the collection, through either microfilm copies or bound paper copies.

The Digitisation of Theses

With the growth of computer usage over the last 20 years we have seen the gradual development of the notion of submitting a thesis or dissertation in digital form into an electronic or digital repository.

¹³⁷ National Library of Australia, *Theses* (2006) <<http://www.nla.gov.au/apps/eresources/action/item?id=1484&loaditem=true>> at 30 June 2006.

¹³⁸ The British Library, *British Thesis Service* (2006) <<http://www.bl.uk/services/document/brittheses.html>> at 30 June 2006.

For example, the Australasian Digital Theses Program (ADT Program) was established in order to improve access to, and enhance the transfer of, research data contained in theses through the provision of full text theses available on the Internet. It establishes a distributed database of digital versions of theses produced by postgraduate students at all Australian universities, which is made available on the Internet. The aim behind the ADT Program is to provide access to, and to promote, Australian research to the international community through the reproduction of theses on the ADT database.

Given that it is the responsibility of each individual institution to maintain an archived copy of the theses, every member of the ADT Program is required to host their own theses on a server located within the university. However, every member uses an identical database configuration, standards and metadata, ensuring compatibility with all electronic theses contained in the ADT Program.

Copyright Management Issues for Electronic Theses and Dissertations

With the increasing trend towards the promulgation of research findings electronically, there has been a concomitant increase in the number of Australian academic institutions that have 'put online' electronic versions of dissertations and theses. Accordingly, there is a need for comprehensive protocols for managing the copyright issues in providing access to Electronic Theses and Dissertations (ETDs).

To build protocols for managing the legal aspects involved in making ETD available online, it is necessary to consider the issues from the perspective of each of the following four distinct stakeholders:

1. **The student.** As the contributor of original material, the submitting student will have IP rights in most, if not all of the content. This will include copyright, but may also have patent issues arising (for example, containment of pre-patent disclosure).¹³⁹

¹³⁹ Publication prior to the filing of a patent will usually result in the inability to get the patent, as the invention would no longer be 'novel'. There are now some provisions for grace periods.

2. **The supervisor.** Depending on the discipline, there may be some content of the thesis that is directly or co-contributed by the student's supervisor. This may give IP rights to the supervisor and/or the supervisor's employer, the relevant academic institution.
3. **University, granting agency and industrial partner.** Universities, granting agencies and industrial partners typically have IP rights agreements and policies that may govern some of the ETD content.
4. **ETD disseminating institution (Repository).** Institutions that have a repository of ETD need clarification of IP rights ownership. What is the status of the repository? (Is it a publisher?); what are the permissions required for cited materials; and are there any exemptions available (such as fair dealing for research or study, or criticism or review)? There may also be tortious issues arising in rare circumstances (such as defamation or passing off).

Adopting the perspective of each of these stakeholders, the management of IP rights in ETD needs to be considered at a fine level of granularity. Taking this approach, numerous questions arise, including:

1. How to manage licensing of distribution?
2. How is the whole work in the thesis and dissertation to be regarded (in other words, is it entirely an original work of the student or does it contain third party or other contributions)?
3. Is this discipline dependent?
4. How to manage cited materials?
5. How to manage contributions by others? (for example, technical photos, cite charts etc.)
6. How to manage derivative works?
7. How to manage confidential information (for example, pre-patent materials)?
8. Liability and risk management?
9. What protocols should be adopted?

The key objective of copyright management in this context is to ensure that the ETD repository has appropriate authorisation to be able to legally carry out all the acts involved in putting the ETD online. In other words, the ETD repository must be granted a licence (preferably in written form) by the copyright owner — usually by the author of the thesis — authorising the ETD repository to reproduce and communicate or otherwise disseminate the thesis via the Internet. Where third party copyright material is included in the ETD, it will be necessary to ensure that appropriate ‘clearances’ (in other words, permissions) have been obtained to use that material in the ETD, unless permission is not required under law.

Status of the Repository – is it a ‘Publisher’?

Copyright issues facing ETD repositories may include whether the repository is a publisher or a ‘re-publisher’ of the thesis for the purposes of copyright, defamation, confidential information (trade secrets) and privacy issues.

In terms of copyright, where a thesis in hard copy form in the form of a work (literary, dramatic, musical or artistic work) is digitised and made available online in an ETD repository where it can be accessed and downloaded by members of the academic and research community, it is arguable that it would be deemed to have been published on the basis of the operation of s 29(1)(a) of the *Copyright Act*. However, the deemed publication provision has a much narrower scope of operation in relation to cinematograph films (s 29(1)(b)) than for ‘works’. Publication is only deemed to occur if copies of the cinematograph film have been sold, hired, or offered or exposed for sale or hire to the public. While it is arguable that copies of film-based ETD are supplied to the public when they are made available for access in an ETD repository, the absence of any commercial dealings in the way of sale, hire, etc. means that it is not possible to rely on the deemed publication provision.¹⁴⁰ Since ETD consisting of moving images (and attracting copyright protection as cinematograph films) will not have the benefit of the deeming provision, it will be necessary to consider whether non-commercial distribution of film ETD from ETD repositories where they

¹⁴⁰ *Copyright Act* s 29 (1)(b).

can be accessed by members of the academic and research community can amount to publication.

Converting Paper Theses to Digital Theses

Where any paper thesis is converted to a digital thesis (p2ETD) a number of copyright issues may arise. These include scanning the thesis without permission of the copyright owner, which will breach copyright as it involves the exercise of the copyright owner's rights of reproduction.¹⁴¹

Furthermore, in retroactively distributing electronic versions of paper based theses (especially older theses) there is the difficulty in getting the permission of the author. Obtaining such permissions would be expensive (both in terms of time and actual fees). One suggested option is to adopt a risk management approach and engage in the digitisation and digital archiving process anyhow given that the risk of copyright infringement proceedings commencing is low.¹⁴²

Another problem with older theses is that even if the author is located, it is unlikely that the author will invest much time or money in establishing that use of any third party content copied is permitted or indeed engage in resolving any of the issues that may arise. Therefore, considerable caution needs to be taken when dealing with the authors of paper based theses and a more specialised licence agreement may be needed.

¹⁴¹ *Copyright Act* ss 31, 101. See also Sections 51 and 53 of the *Copyright Act 1968*. Section 51(2) applies to a manuscript or a reproduction of an unpublished thesis or other similar literary work that is kept in a library or a university or other similar institution. It provides that copyright in the thesis or other work is not infringed by the making or communication of a reproduction of the thesis or other work by or on behalf of the officer in charge of the library if the reproduction is supplied (whether by communication or otherwise) to a person who satisfies an authorized officer of the library that he or she requires the reproduction for the purposes of research or study. Section 53 extends the application of section 51 to illustrations accompanying the thesis or other work. See further, Hudson and Kenyon, *Guidelines for Digitisation*, 129.

¹⁴² See Hudson and Kenyon, *Guidelines for Digitisation*, 50. Arguably, authors of theses would be happy to have their thesis distributed. The greatest risk of copyright infringement would arise if the student assigned the copyright in their thesis to a third party, such as a publisher, and the publisher sought to take action against the repository for breach of their reproduction and communication rights.

Third Party Copyright in Electronic Theses and Dissertations

A high proportion of ETD will contain third party materials, in the form of quotes of text passages, drawings, photographs, reproductions of paintings, video and sound clips and so on. It is essential for ETD repositories to develop and implement strategies to avoid incurring liability (whether through an action for copyright infringement or through a request for payment of equitable remuneration to a copyright collecting society) due to the unauthorised use of any third party copyright materials included in ETD.

If the copyright owner of the third party content has given permission for the work to be used, repositories must ensure that the terms of such permission are not only confined to use in the original theses or dissertation but extend to reproducing or communicating the content for the purposes of digitisation and public access via the repository. The use of third party copyright materials in ETD will typically involve acts within the scope of the copyright owner's exclusive rights to reproduce¹⁴³ or make a copy¹⁴⁴ and to communicate to the public.¹⁴⁵

Exercise of the Reproduction/ Copying Right

Incorporation of third party materials into the new copyright work created by the student, in other words, the ETD, whether in the form of a quote of a passage of text from a literary work, inclusion of a diagram or samples of digital images or sounds, will involve the exercise of the reproduction or copying right. Where an ETD is born digital, it will be the student (rather than the university) who does the initial reproduction and copying of the third party material although the consequences of any further reproduction or copying made by the repository need to be considered. Note that, in the case of a thesis submitted by the student in hard copy form, the reproduction right will be exercised by the university when it converts the work from hard copy into digital format.¹⁴⁶

¹⁴³ For Part III works.

¹⁴⁴ For Part IV subject matter.

¹⁴⁵ 'Communicate' is defined in s 10(1) as meaning to 'make available online or electronically transmit (whether over a path, or a combination of paths, provided by a material substance or otherwise)'.

¹⁴⁶ *Copyright Act* s 21 (1A).

Exercise of the Communication Right

Making an ETD available on a repository website where it can be accessed by users involves an exercise of the communication right, which encompasses making copyright material available online or electronically transmitting it.¹⁴⁷ In a system which is designed so that the ETD is uploaded to the repository directly by the student, it may be that only the student engages in an act of communication. However, in the situation where the student provides the repository with the ETD and authorises the repository to make the ETD available online but all further steps required to make the ETD available online at the repository's web site are carried out by the repository, it is likely to be the case that the act of communication is done by the repository.

The only guidance provided by the *Copyright Act* is found in s 22(6) which states that 'a communication ... is taken to have been made by the person responsible for determining the content of the communication'. The question which arises is whether it is the repository or the student who is the 'person responsible for determining the content of the communication'.¹⁴⁸

Due to the intimate connection the university has with the inception, completion and then uploading of the thesis there is strong argument that it has either undertaken an act of communication or authorised such an act.¹⁴⁹ If the university has undertaken the primary act of infringement (in other words if it actually undertook the infringing act, namely communication) then liability accrues regardless of fault, subject to the exceptions already highlighted. If the university has merely authorised the act of communication then a number of 'fault based' factors will need to be considered including the power to prevent the act, the relationship between the university and the infringer (student) and whether the university took reasonable steps to avoid the act

¹⁴⁷ *Copyright Act* s 10 (1).

¹⁴⁸ See further *Universal Music Australia Pty Ltd v Cooper* [2005] FCA 972, [70]-[76].

¹⁴⁹ For this reason it would seem unlikely that the university could rely on ss 39B, 112E *Copyright Act* which state that merely providing facilities to make or facilitate the making of a communication is not, without more, an authorisation of copyright infringement: *Universal Music Australia Pty Ltd v Cooper* [2005] FCA 972, [97]-[99]; *Universal Music Australia Pty Ltd v Sharman License Holdings Ltd* [2005] FCA 1242, [418]. In relation to moral rights see *Copyright Act* s 195AVB.

(including complying with any industry codes of practice). Regardless of which argument is correct, due to the university's close connection with the thesis, its risk of liability for communicating the thesis must be carefully managed.

There are a number of options available to a repository in order to mitigate the risk of copyright infringement in relation to third party content for born digital theses and dissertations.

These include:

1. Ensuring that ETD candidates are provided with sufficiently extensive information and, if necessary, practical training on the basic principles of copyright law, so they understand when they can use third party content in their thesis without permission (in other words, an insubstantial part or a substantial part which can be used because of the operation of the fair dealing or other exception to infringement) and when they will need to obtain permission ('clearance') from the copyright owner to use third party content and how to obtain permission.
2. Requiring the ETD candidate to be responsible for identifying all third party content included in the thesis, determining which third party content they require permission to use and obtaining all necessary licences (typically a non-exclusive, perpetual licence) from the owners of such third party content, which must be broad enough to permit the thesis containing the third party material to be reproduced and communicated via the Internet (whether by the student, the university or the disciplinary repository).
3. Requiring the ETD candidate to 'self manage' any third party content which is not authorised for digital distribution.

Copyright law does not require permissions where an insubstantial amount of a third party copyright work is involved or where an exception such as fair dealing applies. However the operation of both these doctrines is very fact specific. The best that can be done is to provide ETD candidates with clear examples of what the courts have

decided in the past so they have a practical understanding of what material they can use and when they should seek permission. For example, in *TCN Channel Nine Pty Ltd v Network Ten Pty Ltd (No 2)*¹⁵⁰ it was held that whether a part taken is a substantial part or not, involves an assessment of the importance of the part taken to the work as a whole.¹⁵¹

As discussed in ‘Overview of Principles of Copyright Law’, above, copyright is not infringed by dealings with copyright materials that are considered to be ‘fair’ and provided the dealing falls within one of five classes of purpose:

- research or study (ss 40 and 103C)
- criticism or review (ss 41 and 103A)
- reporting news (ss 42 and 103B)
- judicial proceedings or professional advice (ss 43 and 104)
- parody or satire (ss 41A and 103AA).

Once it is established that the purpose for using the third party copyright material fits into one of these categories, the next step is to consider whether the use made of that material for that purpose is fair.

In the context of ETD, the most relevant of the fair dealing provisions are those which exempt from infringement dealings with copyright materials for the purposes of ‘research or study’ and ‘criticism or review’.¹⁵² The terms ‘research’ or ‘study’ are not defined in the *Copyright Act*, however, in *De Garis v Neville Jeffress Pidler*¹⁵³ Beaumont J held that term ‘research’ within the meaning of s 40 of the *Copyright Act* is intended to have its ordinary dictionary meaning:

25. According to the Macquarie Dictionary, ‘research’ may be defined as

¹⁵⁰ [2005] FCAFC 53 (Sundberg, Finkelstein and Hely JJ, 26 May 2005) [52].

¹⁵¹ *Network Ten Pty Ltd v TCN Channel Nine Pty Ltd* (2004) 78 ALJR 585, 589, 605; *TCN Channel Nine Pty Ltd v Network Ten Pty Ltd (No 2)* [2005] FCAFC 53, [12], [50]–[52]; *Network Ten Pty Ltd v TCN Channel Nine Pty Ltd* [2005] HCA Trans 842 McHugh and Kirby JJ. See also *Haines v Copyright Agency Ltd* (1982) 42 ALR 549.

¹⁵² *Copyright Act* ss 40–43, 103A, 103B, 103C, 104.

¹⁵³ *De Garis v Neville Jeffress Pidler* (1990) 18 IPR 292, 298 (hereinafter *De Garis*).

‘1. diligent and systematic enquiry or investigation into a subject in order to discover facts or principles: research in nuclear physics ...’

Similarly, the *Copyright Act* does not define ‘criticism’ or ‘review’, although it has been held that the words are of ‘wide and indefinite scope which should be interpreted literally’. In *Warner Entertainment Co Ltd v Channel 4 Television Corp PLC*¹⁵⁴ Henry LJ stated that the question to be answered in assessing whether a dealing is fair or not is ‘is the [work] incorporating the infringing material a genuine piece of criticism or review, or is it something else, such as an attempt to dress up the infringement of another’s copyright in the guise of criticism’.

It is clear from the judicial consideration of the meaning of these terms¹⁵⁵ that an individual student engaged in activities involving the use of third party copyright material in the course of researching and writing a thesis would be able to establish that their acts are for the purposes of ‘research or study’ or ‘criticism or review’. It is also clear from the wording of ss 40 and 41 that the fair dealing provisions can be raised as a defence to copyright infringement in relation to an act of communication. Furthermore, there does not seem any doubt that a student can rely on the fair dealing provisions to communicate copyright material for the purposes of ‘research or study’ or ‘criticism or review’. The only doubt raised here is whether any act of communication¹⁵⁶ by the university can be regarded as being for the purposes of ‘research or study’ or ‘criticism or review’?¹⁵⁷ This is explored in the following two arguments.

¹⁵⁴ (1993) 28 IPR 459, 468.

¹⁵⁵ *De Garis*, 298; *CCH Canadian Ltd v Law Society of Upper Canada* [2004] 1 SCR 339; *Warner Entertainment Co Ltd v Channel 4 Television Corp PLC* (1993) 28 IPR 459, 468; *TCN Channel Nine Pty Limited v Network Ten Limited* [2001] FCA 108 [66], also see [16]–[17]; see also *TCN Channel Nine Pty Ltd v Network Ten Pty Limited* [2002] FCAFC 146.

¹⁵⁶ If the university is not regarded as undertaking an act of communication but rather authorising an act of communication then the issue will stand or fall on the basis of the student’s ability to rely on the defence.

¹⁵⁷ See generally: *De Garis v Neville Jeffress Pidler Pty Ltd* (1990) 189 IPR 292; cf. *CCH Canadian Ltd v Law Society of Upper Canada* [2004] 1 SCR 339.

Argument 1: The student's act of research or study or criticism or review includes dissemination of the end product and the university in communicating the ETD is part of that process.

If the university can successfully argue that it is simply a part of or extension of the student's activities and merely a conduit for dissemination then it is more likely that a court will accept an argument that the university in communicating the ETD is doing so for the purpose of research or study or criticism or review. For what other purposes is the university engaging in this activity? Is it to promote the university as a commercial entity or is it to disseminate a product of research or review?

In the old hard copy world the student reproduced copies of the thesis usually through a copying service, supplied them to the university and they were placed on the library shelf. History tells us that no one has ever questioned in the hard copy world the act of the thesis copying service in terms of copyright infringement and the applicability of the fair dealing provisions. If anyone had successfully argued that the thesis copying service could not rely on the fair dealing provisions then the thesis would never have been copied or made available for others to read. It seems odd that a similar activity cannot be undertaken with the same degree of legal certainty in the digital environment, especially when technology neutrality is seen to be a key part of our legal framework.¹⁵⁸

Thus, the university could argue that the student's act of research or study or criticism or review includes dissemination of the end product and the university in communicating the ETD is part of that process. In *CCH Canadian Ltd v Law Society of Upper Canada*¹⁵⁹ the Supreme Court of Canada explained that when library staff made copies of legal materials they did so for the purpose of research 'although the retrieval and photocopying of the legal works are not research in and of themselves, they are necessary conditions of research and thus part of the research process'.¹⁶⁰ Dissemination of research is very much part of the modern

¹⁵⁸ Consider *Electronic Transactions Act 1999* (Cth) ss 11(6), 12(6)

¹⁵⁹ [2004] 1 SCR 339.

¹⁶⁰ [2004] 1 SCR 339, [64]. See also *TCN Channel Nine Pty Ltd v Network Ten Pty Ltd* [2002] FCAFC 146 at [100]-[101]: 'Ten engaged Working Dog Pty Ltd (Working Dog), referred to by the primary judge as "its contracted production team", to produce for it a television programme which would, amongst other things, involve criticism and review and the

research process and the university is merely helping this to happen. As the Supreme Court explained a restrictive interpretation of the fair dealing provisions ‘could result in undue restriction of users’ rights’.¹⁶¹

If this argument for the university based on its function as a conduit for dissemination cannot be sustained then the argument for the operation of the fair dealing provisions must focus solely on the nature of the activity being undertaken by the university. In particular, it is unlikely that the university would be able to avail itself of the fair dealing defence for purposes of criticism or review, although such a purpose may well underlie the student’s use of the third party material. However, it is arguable that current concepts of ‘research’ (and, possibly, ‘study’) are sufficiently broad to encompass the dissemination of research outputs by means such as making ETD available for access through web-based repositories. In a recent report, The British Academy stated the following in relation to the broader meaning of research:

UK law has always provided for exemption from copyright for fair dealing in the course of research. There is, however, no statutory definition of research, or clarity on what differentiates the use of otherwise copyright material in research from its use in private study, or in criticism, or in review. Research involves the production of new ideas, whereas private study might represent only the consideration of existing ones. But this is a fine line indeed, and not one that it would seem appropriate for a publisher, or a court, to draw ... But research without the publication of the results is barely if at all distinguishable from private study, and there is little or no public benefit in the production of new ideas unless they are made publicly available.¹⁶²

reporting of news events. The purpose of Working Dog in the production of these programmes was the purpose of Ten. Consistently with the decisions of the UK Court of Appeal earlier referred to, the “purpose” referred to in ss 103A and 103B is to be ascertained objectively, and it was neither necessary nor appropriate for officers of Ten or of Working Dog to give evidence that they had a sincere belief that he or she was criticising a work or an audio-visual item or reporting news.’

¹⁶¹ ¹⁶¹ [2004] 1 SCR 339 [54].

¹⁶² The British Academy, *Copyright and Research in Humanities and Science* (2006), 9 <<http://www.britac.ac.uk/reports/copyright/index.html>> at 25 September 2006.

In the absence of clarity in either statute or case law, we focus on what we believe the position should be. We consider that the research exemption must extend to the publication of research. The exemption would be largely nugatory and the consequences seriously inimical to scholarship if it did not do so. We also consider that the distinction between non-commercial and commercial research should relate to the purpose of the research, rather than the purpose of the publication of the research.¹⁶³

To restrict the concepts of ‘research’ and ‘study’ to the narrow range of activities associated with collecting, reading, summarising and extracting parts of the material may unjustifiably limit the operation of this fair dealing provision. In the digital, networked environment in which research and study now occur and in which research and teaching processes are iterative and collaborative, communicating research findings to an online audience of colleagues and commentators is considered an integral part of the research and teaching process.

Argument 2: The University in communicating the ETD is engaged in an act of research broadly defined as an intermediary.

As explained, the University is either engaged in the act of communicating the ETD or assisting such communication. Amendments to the *Copyright Act* introduced as a result of the AUSFTA limit the liability (by way of limiting remedies available) for certain acts performed by intermediaries.¹⁶⁴ These provisions apply to ‘carriage service providers’ and provide for a ‘safe harbour’ from liability in defined circumstances. They are commonly called the ‘ISP safe harbour provisions’ and are modelled on similar provisions in the US *Digital Millennium Copyright Act 1998* (DMCA). These new provisions limit the remedies available against carriage service providers for copyright infringements that occur on their systems, as long as they comply with certain conditions.

There is currently some uncertainty as to whether universities may take advantage of this scheme. This uncertainty relates primarily to whether

¹⁶³ Ibid 10.

¹⁶⁴ *Copyright Act* ss 116AA-AJ Part V Division 2AA.

universities fall within the definition of ‘carriage service provider’, which for the purpose of the safe harbour provisions is drawn from the highly technical definition provided by the *Telecommunications Act 1997*. From 1997 to 2001 a determination by the then Minister for Communications, Information Technology and the Arts, Mr Richard Alston, under s 95 of the *Telecommunications Act* effectively excluded universities from being carriage service providers by stating that services provided by tertiary education institutions in connection with their research, educational and administrative functions were not carriage services. Since this determination was allowed to lapse, the general opinion seems to be that universities are nevertheless excluded from being carriage service providers because they do not provide their services ‘to the public’, as required by s 88 of the *Telecommunications Act*.¹⁶⁵ In late 2005 the Attorney General’s Department commenced a review of the scope of the safe harbour scheme which, among other things, sought comments on this issue. In making a submission to that Review the AVCC explained:

As the regime currently stands, only carriage service providers (within the meaning of the *Telecommunications Act*) can obtain the protection of the safe harbour regime. As most universities are not engaged in supplying a carriage service to the public but rather to their immediate circle (as that term applies under the *Telecommunications Act*) they do not qualify to take advantage of the safe harbour regime.¹⁶⁶

The Government is yet to announce the findings of this review. Until that point in time we must assume that Universities even if they could

¹⁶⁵ See generally: DEST, *Limitation on Remedies Available against Carriage Service Providers under Part V Division 2AA of the Copyright Act*, submission by the Department of Education, Science and Training to the Attorney General’s Review, October 2005.

¹⁶⁶ AVCC, *Safe Harbour Regime: Review of the Scope of Part V Division 2AA of the Copyright Act*, submission by the AVCC to the Attorney General’s Review, October 2005, <http://www.avcc.edu.au/documents/publications/policy/submissions/AVCC-SafeHarbourSubmission_OCT05.pdf>. It seems commonly accepted that the University of Queensland is the only Australian university that currently falls within the definition of ‘carriage service provider’: AVCC, *University IT Systems: Managing Liability for Transmitting, Caching, Hosting and Linking to Copyright Material* (2004) 2 <http://www.flinders.edu.au/isd/copyright/AVCC_resource_paper.pdf>.

satisfy the condition for enlivening the safe harbour provisions cannot take advantage of them because they are not carriage service providers.

One other suggestion is that any potential infringement of third party content by the University in the ETD process could be covered by the statutory educational licences which allow certain acts on the basis of equitable remuneration. Whether this is the case raises a number of difficult legal questions which deserve closer consideration. However, it is important to keep in mind that the statutory licences do not require remuneration where there is a fair dealing or use of an insubstantial part and as such close scrutiny of the material is the sensible starting point.

Protocols for the practical handling of ETD¹⁶⁷

In light of the foregoing analysis it is clear that universities are subject to the risk of copyright liability for the communication of ETD and as such need to put in place workable and effective compliance mechanism. The sensible way to approach these steps is to have the ETD candidate self manage the process from the very first day of their candidature. That is, the student would be asked to record all third party copyright materials included in the thesis, to make an assessment of the copyright status of these materials and to note this in their Copyright Compliance Table on a continuous basis. In managing these situations the following steps are suggested:

¹⁶⁷ See further, Damien O'Brien and Dr Anne Fitzgerald, *Copyright Guide for Research Students: What you need to know about copyright before depositing your electronic thesis in an online repository*, August 2007, OAK Law Project, including Copyright Compliance Table and Model Third Party Copyright Permission Requests available at <http://www.oaklaw.qut.edu.au/files/Copyright%20Guide%20for%20Research%20Students.pdf>.

1. **Identify all third party copyright materials** included in the ETD.
2. **Is there a substantial part?** Examine each item of third party copyright content included in the ETD to assess if its inclusion involves the exercise of acts (for example, reproduction, adaptation) in relation to a substantial part of the third party copyright content; where only an insubstantial part of any item of third party content is used, there is no need to take further steps as use of an insubstantial part is not an infringement and does not need to be authorised by the copyright owner. Establishing guidelines for what is a substantial part is integral to the risk management process. It is not possible to provide absolute and firm guidelines for all situations, but it must be understood that any figures stated in the guidelines will essentially become the de facto rule.
3. **Is there a fair dealing?** If a substantial part of an item of third party copyright content is included in the ETD, consider whether use of that part is justified under one or more of the fair dealing provisions.
4. **Does any other exception to copyright infringement apply?** For example it is not an infringement of copyright to take a photo of a sculpture or work of artistic craftsmanship that is on permanent public display,¹⁶⁸ so if a student includes an image of such a work in a public place there is no need to obtain permission from the owner of copyright in the publicly displayed work. A list of these kinds of miscellaneous exceptions which are relevant to the education sector should be compiled. A dated but useful starting point for understanding these exceptions is found in the Copyright Law Review Committee's (CLRC) reports, *Simplification of Copyright Act 1968 Part 1* (1998)¹⁶⁹ and *Copyright and Contract* (2002).¹⁷⁰

¹⁶⁸ *Copyright Act* s 65.

¹⁶⁹ See: <http://www.clrc.gov.au/agd/WWW/clrHome.nsf/Page/Overview_Reports_Simplification_of_the_Copyright_Act_Part_1>.

¹⁷⁰ See <http://www.clrc.gov.au/agd/WWW/clrHome.nsf/Page/Overview_Reports_Copyright_and_Contract>.

5. **Permission requested?** If after going through these steps there is still uncertainty about whether the use of the third party content in the thesis is authorised, a request should be sent to the copyright owner specifying the third party materials which are to be included in the thesis and the use to be made of that material and seeking express permission for such use; any licence obtained for the use of third party content must be broad enough to permit the thesis to be reproduced in digital form and communicated online (whether by the student, the university or a disciplinary repository). Since there will be doubt about whether the reproduction and communication of some materials included in theses is permissible, in some cases there will be no option but to seek express permission.

Adopting appropriate licences

In general, repositories will be seeking to rely on non-exclusive licences from owners of copyright in theses which they seek to place in the repository. The four types of licences listed below should be considered in relation to licensing issues for ETD.

Deposit licence – between the owner of copyright in the ETD and the ETD repository in order to give certainty to repositories, in terms of what rights they have to store, manage and organise the ETD stored within the repository. The licence could also contain terms that reduce repository liability through disclaimers and indemnities.¹⁷¹

End user licence – the end user, in other words, the person who downloads a thesis, should be clearly informed about the specific activities of use and re-use that are permitted — under what is termed an End User Licence. For example, this would typically include activities such as browsing (reading on screen); downloading and printing; or possibly downloading and distributing copies in class. To ensure that end users are clearly informed of the uses they are permitted to make of ETD, it is

¹⁷¹ Examples of current deposit licences include Swinburne University of Technology *Access to Thesis* <<http://www.swin.edu.au/research/postgrad.htm>> and the National Library of Canada *Theses Non-Exclusive License* <<http://www.nlc-bnc.ca/obj/s4/f2/frm-nl59-2.pdf>>.

recommended that a standard, though flexible, protocol be adopted for end user licensing. For example, a straightforward approach would be for the ETD holder to license end users under one of the standard open content licences such as the Creative Commons (CC)¹⁷² or AEShareNet licence.¹⁷³

Third party licence – as explained at length above, where third party copyright content is included in the ETD it is necessary to confirm that rights to use the content have been granted by the third party copyright owner (in the absence of any exemption or exclusion from copyright infringement).¹⁷⁴

Publisher licence – a licence between the publisher and the ETD repository will be crucial where an ETD candidate has already assigned the copyright in all or part of their thesis, such as where they have had an article published prior to submitting the electronic thesis and dissertation.

CONCLUSION

The challenge for knowledge management lies in harnessing the enormous power of networked digital technologies. At the heart of this issue is best practice copyright management. What we have shown in this article is that to achieve this, institutions and people have to appreciate the variety of copyright management models that are emerging and how to employ them. If open access is a value we wish to promote, for social, economic and cultural reasons, institutions must articulate their commitment in clear policies. From this touchstone an effective copyright management framework can be built. At the end of the day, we must realise that better copyright management will provide us with more choices (including open access) but that it will not happen by default. It must be structured and managed. That is the challenge and the path forward.

¹⁷² See by way of example: Oleg Evnin's CC licenced Caltech doctoral thesis at <<http://resolver.caltech.edu/caltechetd:etd-06072006-174745>> at 13 July 2006.

¹⁷³ See <<http://www.aesharenet.com.au>>.

¹⁷⁴ For a current example see Queensland University of Technology, *Non-exclusive Licence Agreement for Inclusion of Third Party Copyright Material in ETD* <http://www.research.qut.edu.au/downloads/ADT_copyright_owner_request.doc>.

PART FIVE

CONTRACTUAL AND
TECHNOLOGICAL
FRAMEWORKS

THE UNIVERSITY–INDUSTRY DEMONSTRATION PARTNERSHIP: AN INCREMENTAL IMPROVEMENT TO UNIVERSITY-INDUSTRY COLLABORATION

James J Casey, Jr, Esq¹

INTRODUCTION

The cover of the July/August 2007 issue of the *Harvard Business Review* has two phrases that sum up the goals of university-industry collaboration: ‘Managing for the Long Term’ and ‘Going the Distance’. Although those phrases were meant for companies, these phrases accurately reflect what university-industry collaborations and the University-Industry Demonstration Partnership (UIDP) are all about.

¹ Director of Contracts and Industrial Agreements, University of Texas at San Antonio.

All comments herein are solely my own and do not represent the opinions or positions of the University-Industry Demonstration Partnership. Special thanks are extended to Robert Killoren, UIDP President, Dr Susan Butts, UIDP Vice-President, and Dr Merrilea Mayo, GUIRR Director, for their assistance and counsel. Dr Casey has previously written on university-industry collaborations. See ‘Long-Term University/Industry Collaborations’ (June 2007) *R&D Magazine Academic Sourcebook* 26–7; ‘University-Industry Connections: A Small School Perspective’ (June 2006) *R&D Magazine Academic Sourcebook* 9, 11; ‘Developing Harmonious University-Industry Partnerships’ (2004) 30 *Dayton Law Review* 245–63; ‘Enhance University-Industry Collaboration’ (2005) *Milwaukee Business Journal*, June 3, A53; ‘Making a Good Thing Even Better’ (Fall 2004) 14(2) *Research Management Review* 10–22. In addition, Dr Casey was co-editor and contributor to *Living Studies in University-Industry Negotiations: Applications of the Guiding Principles for University-Industry Endeavors* (National Council of University Research Administrators and the Industrial Research Institute, April 2006). Additional articles by Dr Casey that may be of interest include: ‘An Era of Uncertainty, An Era of Opportunity’ (July/August 2007) *NCURA Newsletter* 4–5; ‘Giving it Away: Free Technology Transfer to the SME Sector’ (Spring 2006) 15(1) *Research Management Review* (with Dr Peter Kavanagh and Mr Andy Maguire, Dublin Institute of Technology); ‘The Legal Dimensions of Research Administration’ (Winter 1998) 10(1) *Research Management Review* 7–17.

Put another way, university-industry collaborations are critical long-term infrastructure developments. Managing for the long term should be a goal for all managers of these partnerships.

This chapter addresses a number of topics. First, it discusses recent initiatives in the United States to strengthen these collaborations, particularly the University-Industry Partnership Project (UIPP). Then, it transitions to discuss the UIDP and its first demonstration project—TurboNegotiator (TN). Lastly, the chapter makes some concluding observations about the UIDP, TN, and university-industry collaborations in general.

Background

University-Industry collaboration is a critical topic currently being discussed in US academic, industry, and government circles. With federal research funding being in a state of zero growth or actual decline (depending how you want to slice the numbers), colleges and universities are being forced to diversify their sponsored program and research portfolios. This includes looking for new funding sources from private business, corporate foundations, and other non-profit foundations. In addition, the new and expensive costs of US homeland security and the wars in Afghanistan (2001) and Iraq (2003) are having a significant fiscal impact upon the future shape of the US budget. I noted these facts in the article I wrote for the July/August issue of the *NCURA Newsletter* (National Council of University Research Administrators), and made the further observation that research/R&D funding in the United States does not occur in a vacuum.² It is intimately tied to other policy choices that President Bush and the Congress make. These same choices also occur in other western democratic societies.

Another important dimension is how industry is investing its R&D funds. As an article in the July/August 2007 issue of the *Harvard Business Review* points out, from 2000–07 R&D has focused on new projects rather than on directed basic research.³ As author Jim Scinta, chair of the Industrial Research Institute's Research-on-Research Committee

² James J Casey Jr, 'An Era of Opportunity', (2007) July/August *NCURA Newsletter*, 5 <<http://www.ncura.edu/content/misc/newsletter/news724.pdf>>.

³ Jim Scinta, 'Where More R&D Dollars Should Go' (July/August 2007) *Harvard Business Review* 26.

points out, firms that dedicate a disproportionate amount of their R&D on new projects rather than basic research will probably satisfy some of their intermediate business goals but will fail to cultivate the broad-based knowledge that will ultimately lead to long-term growth through innovation.⁴ While the article does not specifically address business collaboration with higher education, a decline in basic research will most probably impact higher education research facilities that are capable of engaging in basic research.

Furthermore, over the past few decades American universities have enjoyed a strong productive relationship with private companies. On a general level there seems to have been a broad consensus that these relationships are important to the United States, domestically and internationally. Both sides have, by and large, found these relationships to meet their mission requirements. There is much to celebrate in this recent history of collaboration.

WHY UNIVERSITIES AND INDUSTRY COLLABORATE

Why do universities and industry collaborate? As I have previously written, the reasons are many, though this list is by no means exhaustive:⁵

- Universities provide a *ready pool of graduate and undergraduate students* that industry may access for their work requirements. Students in return receive critical workforce training that supplements coursework. Workforce training is increasingly recognised within the US as a critical component of education in a knowledge-based, international economy.
- *Technical opportunities exist in industry* for faculty and students that may not exist in institutions of higher education.

⁴ Jim Scinta, 'Where More R&D Dollars Should Go' (July/August 2007) *Harvard Business Review* 26.

⁵ James J Casey, Jr, 'Developing Harmonious University-Industry Partnerships' (2004) 30 *Dayton Law Review* 251–2. See also NCURA, GUIRR & IRI, *Industry-University Focus Group, National Council of University Research Administrators Annual Meeting* (2003).

- *Materials exist in industry* for research and educational purposes that may not exist in institutions of higher education.
- Collaborations with industry provide *research funding to universities*, a need that has become increasingly apparent over the past 10 years. Universities come to rely on the generation of extramural funding as they structure their budgets. A sad reality, though, as money should not drive every decision made within universities.
- Such collaborations can advance the *service mission* of universities, an increasing component of universities as they become more involved in their local communities. Such service has also been demanded by local and state governments within which the institutions are located; this could be considered a quid pro quo for tax-exempt status-or at least to forestall political retaliation against universities that are perceived to be 'rich islands' within some communities.
- Collaborations provide for *local and regional economic development*. There is evidence to suggest that university-industry collaborations contribute to the overall economic development of the United States. This is necessary in a post-industrial, knowledge-based economy.
- Collaborations between universities and industry often are *novel to high technology areas*, as opposed to low technology areas (such as basic manufacturing). Nanotechnology and materials science/engineering are examples of such high technology fields. However, the argument is being increasingly made that basic manufacturing is now 'high technology' and hence is important to the overall US economy.
- At some universities these collaborations are part of their *internal reward structure* (financial incentives to faculty which are critical for research development and retention of 'star' faculty). If universities seek to increase their research and sponsored program

portfolios, they must create reward structures for faculty and staff that bring in such extramural funding.

- *Universities often have research infrastructure that industry wants.* For many companies, it is simply more cost effective to contract out research to universities that have the research infrastructure in place rather than building from the ground up or renovating existing facilities.
- *Collaboration is encouraged by the US Government.* Whether through such legislation as the *Bayh-Dole Act of 1980* or through specific programs such as the NSF Partnerships for Innovation Program (NSF-PFI), the US Government explicitly encourages such collaborations. The NSF-PFI Program is an excellent example of combining intellectual property, workforce development, and R&D components into a consistent funding program.
- *Industry outsourcing to universities,* to reduce the costs of doing business and increase profits.

As this list illustrates, this symbiotic relationship reflects benefits to each partner. This is one strong characteristic of university-industry collaboration.

However, there have been recent indications that this relationship is strained and needs some tending to, primarily though not exclusively related to issues of contract negotiation and intellectual property.⁶ According to recent statistics from the American Association for the Advancement of Science (AAAS), there has been a recent decline in the

⁶ For additional background information on university-industry collaboration, the *Bayh-Dole Act of 1980*, the positives and negatives of *Bayh-Dole*, and suggested improvements in the law and regulations to make *Bayh-Dole* even more effective, see the testimony of Dr Susan B Butts, Senior Director, External Science and Technology Programs for the Dow Chemical Company. Her testimony of July 17, 2007, given before the US House of Representatives Committee on Science and Technology, Subcommittee on Technology and Innovation, recommends small changes in *Bayh-Dole* and tax regulations to clarify the intent of Congress relative to ownership/control of IP resulting from industry-sponsored research, with the intent to improve the climate for university-industry research partnerships in the United States. Testimony of Dr Susan B Butts, 17 July 2007, 1. She also reiterated an issue that surfaced during the University-Industry Partnership Project; namely, that foreign universities are more flexible with IP ownership and control, causing more sponsored research to be conducted abroad. Testimony of Dr Susan B Butts, 17 July 2007, 1.

level of industry support for university research in the United States. This information can also be found in the article I wrote for the *NCURA Newsletter* (July/August 2007).⁷ There is also evidence that suggests foreign universities are conducting increasing amounts of US industry sponsored research because they are willing to forgo ownership of intellectual property resulting from the research. American universities are more likely to demand sole ownership of IP generated from university research than their foreign counterparts.

As a September 2006 NSF InfoBrief stated: ‘A three-decades-long trend of increasingly strong ties between industry and universities may have ended.’ This sentiment is confirmed by the AAAS statistics provided for the *NCURA Newsletter* article.⁸

All of these factors lead to the conclusion that, despite a strong historical relationship between US higher education and industry, the present time is an uncertain era for these collaborations. Thankfully, it has been recognised by universities, companies, and the US Government that this trend must be reversed by more vigorous and successful collaborations.

PRIOR EFFORTS BY THE GOVERNMENT- UNIVERSITY-INDUSTRY RESEARCH ROUNDTABLE (GUIRR) AND THE INDUSTRIAL RESEARCH INSTITUTE (IRI)

From the 1980s until the convening of the University-Industry Congress in 2003, the Government-University-Industry Research Roundtable (GUIRR, part of the National Academies in Washington, D.C.) and the Industrial Research Institute, Inc. (IRI) were concerned with strengthening and improving university-industry partnerships.

GUIRR was created in 1984 in response to the report on the National Commission on Research, which called for an institutional forum to facilitate dialog among the top leaders of government and non-government research organisations. GUIRR’s formal mission was revised in 1995 to ‘convene senior-most representatives from

⁷ James J Casey Jr, ‘An Era of Opportunity’, (2007) July/August *NCURA Newsletter*, 5 <<http://www.ncura.edu/content/misc/newsletter/news724.pdf>>.

⁸ James J Casey Jr, ‘An Era of Opportunity’, (2007) July/August *NCURA Newsletter*, 5 <<http://www.ncura.edu/content/misc/newsletter/news724.pdf>>.

government, universities, and industry to define and explore critical issues related to the national and global science and technology agenda that are of shared interest; to frame the next critical question stemming from current debate and analysis; and to incubate activities of on-going value to the stakeholders. This forum will be designed to facilitate candid dialogue among participants, to foster self-implementing activities, and, where appropriate, to carry awareness of consequences to the wider public.⁹

The IRI is the foremost business association of leaders in R&D working together to enhance the effectiveness of technological innovation in industry.¹⁰ Founded in 1938 through the National Research Council, the IRI is comprised of senior executives from a diverse range of industries whose member companies are investing \$70 billion annually in R&D projects worldwide. The IRI is the only cross-industry organisation providing the R&D community with insights, solutions, and best practices in innovation management developed through collaborative knowledge creation.

These efforts were primarily concerned with the creation and modification of a variety of standard/boilerplate contractual agreements, including research agreements. Publications were released for university and private sector use, and over the years these model and boilerplate agreements became part of the university-industry partnership culture. There is no doubt that these model and boilerplate agreements served their roles well, and helped advance the growth of these partnerships. However, it is generally recognised that these publications only addressed part of the relationship, and certainly didn't have a profound impact on improving and managing these relationships in the more dynamic, internet-driven world of the 1990s and the present decade. As a result, it was generally recognised by GUIRR and IRI that more needed to be done within the relationship than promulgate new contract templates. The next section discusses their next move, partnering with NCURA (National Council of University Research Administrators).

⁹ GUIRR 2002 *Annual Report* 2.

¹⁰ Industrial Research Institute, Inc <<http://www.iriinc.org/>>.

UNIVERSITY-INDUSTRY CONGRESS/UNIVERSITY-INDUSTRY PARTNERSHIP PROJECT (UIPP)

The University-Industry Congress was established in 2003 by NCURA and the IRI, with GUIRR serving as the neutral convener. Subsequently the University-Industry Congress was renamed the University-Industry Partnership Project (UIPP).¹¹ The UIPP existed from August, 2003 through its national summit in April, 2006.

NCURA, founded in 1959, is a professional organisation of individuals with interests in the administration of sponsored programs (research, education, and training) primarily at colleges and universities. With over 6000 members nationally and internationally, NCURA serves its members and advances the field of research administration through cutting edge professional development programs.¹²

The UIPP brought together approximately 35 hand picked people from academia, industry, and the US Government. A significant strength was the wide breath of participants, representing small and large universities, small and large companies, and different sectors of the US economy. I was chosen as a delegate because of my legal expertise and experience working at research and non-research universities. The purpose was to discuss the university-industry relationship, ascertain what was working and what was not, and to establish deliverables which would strengthen these relationships.

In broad terms, the UIPP was focused on the following: 1) Turning challenges into successes. This means surmounting the primary challenges of contract negotiations and intellectual property into positive results; 2) Building trust and teamwork. In the first year of the project, it was readily apparent that there was significant distrust among some of the participants, either on a general level or based upon prior bad

¹¹ For additional information regarding the UIPP and UIDP, please see the appropriate sections of the *GUIRR 2006 Annual Report*. This report provides sections on the following: 1) Deemed Exports: Promoting Change on Critical National Issues; 2) *The Here or There?* Report: Bringing New Knowledge to the Debate Over Corporate R&D Globalization; 3) The University-Industry Partnership: An Action Agenda for More Effective Cooperation; 4) UIDP: A New Institution to Strengthen the US Research Enterprise; 5) Major Workshops and Convocations: Advancing National Science and Technology Policy; and 6) The Federal Demonstration Partnership (FDP): A Track Record of Success in Raising Research Productivity.

¹² National Council of University Research Administrators <<http://www.ncura.edu/>>.

experiences; 3) Defining and prioritising the issues. The participants needed to ascertain what the major problem areas were and deciding which ones needed to be addressed first; 4) Finding a ‘common cause’. This is primarily based upon building trust, finding common areas of concern and redress, and creating an action plan; 5) Developing flexibility. This is a recognition that university-industry partnerships, to be truly productive and long-term, must be flexible to meet future demands and changes of an internal and external nature; and 6) Building on existing efforts, such as the Business-Higher Education Forum (BHEF) publication ‘Working Together, Creating Knowledge’ and the Responsible Partnership Initiative by EARMA (European Association of Research Managers & Administrators).

A primary conclusion of this project was that negotiation of sponsored research agreements is a barrier to industry-university research collaboration in the United States. This barrier is exemplified by longer contract negotiation times, contentious negotiation processes, increasing costs resulting from the increase in length and contention, and little or no benefits resulting from the conclusion of the contract negotiation. This conclusion is not surprising given the efforts dedicated to the issue prior to the establishment of the UIPP. TurboNegotiator (TN)—the first demonstration project of the UIDP—is meant to start addressing this problematic area.

Another significant benefit from the UIPP was greater communication and understanding between the project participants and the institutions/sectors they represented. Whereas the first year represented significant distrust and strained conversations, by the last year significant progress was being made and the communication was more open and solution-focused. Calling this transformational change is not an overstatement.

By the time the UIPP ended with a national summit on April 23, 2006, the UIPP came up with two publications that reflected project deliverables: 1) *Guiding Principles for University-Industry Endeavors*, which articulated a preamble and guiding principles for such collaborations; and 2) *Living Studies in University-Industry Negotiations*, which illustrated a variety of successful and problematic partnerships. This document is cross-sectoral and represented a variety of private sector, university, and government actors. The ‘Living Studies’ publication also mapped perfectly with the ‘Guiding Principles’ publication, illustrating the

guiding principles in action. It is true that the *Living Studies* publication is primarily historical in nature. But it is equally true that these studies are meant to be learned from and applied to the present. That is the essence of what a 'living document' is.

UNIVERSITY-INDUSTRY DEMONSTRATION PARTNERSHIP (UIDP) AND TURBONEGOTIATOR (TN)

The conclusion of the UIPP at the national summit occurred simultaneously with the kick off of the UIDP as its successor project. Membership in the UIDP is dues-based, drawing on the idea that institutions that pay to belong within it have a vested interest to make the UIDP succeed. A membership drive for the UIDP started even before the national summit closing out the work of the UIPP. The membership drive, so to speak, continues to this day. The UIDP had its first meeting in December, 2006 and meets every quarter.

The UIDP was modelled after the Federal Demonstration Partnership (FDP), which began as the Florida Demonstration Project in 1986. The FDP is an association of federal agencies, academic research institutions with administrative, faculty and technical representation, and research policy organisations that work to streamline the administration of federally sponsored research.¹³ FDP members of all sectors cooperate in identifying, testing, and implementing new, more effective ways of managing more than \$15 Billion in federal research grants. The goal of improving the productivity of research without compromising its stewardship has benefits for the entire nation.¹⁴

Now 21 years old, the FDP is widely accepted as a success by universities and the federal government as being a model to drive institutional change on a national level. FDP continues to move forward, seeking to improve institutional stewardship of federal research money while ensuring the timely and expeditious conduct of research.

Taking the UIPP results into a concrete realm, and using over 20 years of FDP experience, the mission of the UIDP is to nourish and expand

¹³ The Federal Demonstration Partnership, *About FDP*
<http://www.uidp.org/ABOUT_UIDP.html>.

¹⁴ The Federal Demonstration Partnership, *About FDP*
<http://www.uidp.org/ABOUT_UIDP.html>.

collaborative partnerships between universities and industry in the United States. How will this mission be accomplished? The UIDP states:

The UIDP accomplishes this mission via a coalition of universities and companies who engage in voluntary collaborative experiments or new approaches to sponsored research, licensing arrangements, and the broader strategic elements of a healthy, long-term university-industry relationship. Institutional experiments are chosen and jointly pursued by willing members when they have the potential to increase the level, degree, or ease of university-industry collaboration. A primary focus for the UIDP's initial work will be on streamlining intellectual property negotiations.¹⁵

Forty-nine universities and 20 companies comprise the UIDP as of 6 August 2007. The National Science Foundation (NSF) is a Founder's Circle Member, a category reserved for entities that make a substantial resource contribution to the UIDP. Other members in this category include Pfizer, Ex One, Hewlett Packard, the Kauffman Foundation, and the University of California-Los Angeles.

Benefits of the UIDP

The potential benefits of the UIDP include:

- Improve the research relationship between universities and industry (the focus right now is not on licensing existing university technology funded by the federal government).
- Attract more industry investment into American universities.
- Improve American innovation and competitiveness in a knowledge-based global economy.
- Delivering solutions, not just talk.

¹⁵ University-Industry Demonstration Partnership, *About UIDP*
<http://www.uidp.org/ABOUT_UIDP.html>.

UIDP Characteristics

The UIDP focuses on *collaborative* beta-testing of new approaches to sponsored research, licensing arrangements, and strategic university-industry partnerships. Working groups will be focused on designing institutional experiments.

In addition to these practical, project-related initiatives, UIDP is a forum for the wide dissemination of the latest news, best practices, etc in the area of university-industry collaboration. Institutions that join the UIDP not only belong to demonstration projects; they are part of a broader forum designed to enhance collaboration.

One of the unique characteristics about the UIDP is that it requires a paradigm shift. Whereas the current/past paradigm is characterised by *policy-based contract negotiation* (e.g., the partners have IP policies that drive terms and conditions in agreements), the new paradigm requires a *principle-based paradigm*, one that is characterised by the partners determining the parameters that should be considered in selecting appropriate contract terms and conditions.

It was recognised during the UIPP and now the UIDP that contract negotiations must be conducted in a smarter manner. The knowledge of contract negotiators must be increased. To this end, contract negotiators should know more about the proposed project than just a written statement of work. Here are some of the questions that they should have the answers to:

- Why do the researchers want to work together?
- Who framed the problem that led to the proposed project?
- Who made the creative contributions to the statement of work?
- Who has Background IP that could have an impact on the proposed project?
- Who has key information or materials or prior research results needed for the project to happen?

In the end, though, the proposed contract terms and conditions should be appropriate for the facts of the situation. This illustrates another important theme of the UIDP and a lesson from the UIPP: there is/are no simple template-derived solution/solutions for these partnerships.

Remember, the template-focus best characterised the IRI/GUIRR efforts in the 1980s and 1990s.

Contract negotiators should also know the project parameters, including the following:

- Who had the idea for the research project (professor, sponsor, both simultaneously)?
- Who contributed background technology and background IP?
- Type and importance of non-financial contributions from sponsor (proprietary information, non-commercial materials, results from in-house research, etc.).
- Type and importance of non-labour contributions from the university (specialised equipment/facilities, building on prior research results, etc.).
- Nature of research (fundamental to applied, along a continuum).
- The scientific discipline(s) involved (biology, chemistry, biomedical engineering, civil engineering, etc.).
- The likelihood and/or expectation of inventions resulting from the proposed project.

In my professional career, I *always* endeavoured to learn as much as I could about a potential partnership in advance of negotiating an agreement. Perhaps that was my training as an attorney—negotiating without that information seemed to be negligence.

TURBONEGOTIATOR

The first UIDP demonstration project is TurboNegotiator, a tool to allow university and industry negotiators to rapidly navigate towards mutual agreement on intellectual property provisions (see the main UIDP web page, above). This initiative came out of the UIPP; the latter found that research agreements and intellectual property provisions were among the most significant impediments to past, present, and future collaboration between universities and companies. TN is currently in a

conceptualisation phase (Phase I). Beta testing of TN is at least a year away (late 2008 if not 2009).

The following steps give the reader a strong idea about how TN will work:

- Define and describe the 'Project Space'.
- Populate Project Space with examples of suitable agreement terms.
- Develop a questionnaire to probe parameters for the proposed project and use the answers to map the project into the corresponding sector in Project Space.
- Develop software that will guide the process further. This includes:
 - asking questions based upon input provided by project participants;
 - using responses to map project to a sector within the Project Space; and
 - providing sample agreement terms for that sector (which may include explanations and the positives/negatives for such terminology choices).

As the reader reads on, he or she will see that TN, in theory and in practice, is a multifaceted tool.

TN is a rational basis for building an agreement that accurately reflects the project parameters and what the partners want. It uses example terms as the starting place for negotiations. TN is also a process rather than a solution; it improves understanding of needs and contributions. More importantly, it is an educational process from which all contract negotiators will benefit.

TN is interactive. It will encourage discussion and include input from all key stakeholders. All relevant parties to the agreement should answer the questions. This includes faculty, company researchers, and contract negotiators from all sides involved in the negotiation.

TN is constructive. It suggests terms that are fair and reasonable, and results in less time for negotiation. Projects commence earlier, which is in all everyone's interests. TN will include a time-to-agreement metric,

similar to a tickler file but more sophisticated. This latter component can be an excellent managerial tool to spur quality and time improvements. Quality and speed of negotiation should be the goals and passion of all contract negotiators, regardless of institutional affiliation.

WHAT TURBONEGOTIATOR IS NOT

TN is not a *proscriptive* tool. It does not provide the *right* answer or the *only* answer. *It is not coercive.* If either party is not happy with the outcome, the parties can always walk away from the negotiation or take a different approach or attitude. Maybe the parties have not answered the questions honestly or completely-though this is critical for TN success.

TN does not force or mandate a win-lose outcome. TN seeks to forge agreements that result in productive research, meets the missions of the parties, and possibly lead to long-term relationships. In the end, TN seeks to foster mission compatibility on a project-by-project basis with the desirable outcome of spurring greater thought towards future collaboration.

SUMMARY FEATURES OF TURBONEGOTIATOR

These are the major summary features of TN, given the current status of the project:

- TN has the ability to quickly craft an individualised agreement that allows the research to move forward while meeting the mission needs of each party. Remember the signed contract is a means to the end (the research). It is not the ultimate end, in and of itself.
- TN accepts that contract negotiators are under-trained; hence the focus on TN as being a hands-on resource and tool. Education is critical to research administrators and corporate negotiators alike. While it is commonly assumed that delays to contract completion are due to delays on the university side, the author has found during his professional career that delays also happen as frequently on the corporate side.

- TN requires parties to agree on project scope before proceeding to clause selection. While this sounds like common sense (and is), it seems harder in reality. As a contract negotiator, I always nailed down the project scope before negotiating terms and conditions. As an attorney, it always seemed to me to be negligent to negotiate in the absence of necessary technical/project information.
- TN measures its own success by a 'time to agreement' module. As mentioned earlier, timeliness—along with quality—are the paramount goals of contract negotiation. And there are some areas of contract negotiation, like clinical trial agreements, where time is of the essence. When solving or mitigating medical ailments, not to mention the human subject protocol dimension, it behoves the contract negotiators on all side to reach agreement quickly so that the medical research can go forward.

CONCLUSION

What then can be concluded from TN and the broader forum that is the UIDP? One major point is that both represent incremental improvements to the university-industry partnership in general. There have been efforts in the past to improve this relationship, but they never seemed to permanently provide a forum for ongoing discussion and problem resolution. Good intentions are half the battle, but concrete steps are more important. The world of collaboration is changing rapidly, much more rapidly than in the 1980s and 1990s. Economic, political, and socio-cultural change impacts these relationships. It was logically necessary that a permanent forum be built to tend to this relationship. This is where the UIDP comes in.

Second, both represent a solution-based, incremental focus on university-industry collaboration. There is a time for talk and a time for action. TN in particular represents the action component though the more discussant-focused nature of UIDP. The forum component is equally necessary in a broader context.

Third, TN needs more development but represents a strong move forward. As has been stated earlier, TN is initially designed to handle two-party agreements representing discrete research projects. It is not

initially designed to handle multi-party negotiations nor umbrella projects or master agreements. Obviously, continued development of TN in the latter areas is advisable as those are significant areas within the overall relationship. Also, has been recognised, all sectors of private business need to be represented in the UIDP to be particularly effective.

Another major conclusion is that education and training remain important to professionals, particularly contract negotiators, in both sectors. Skill levels of contract negotiators in both sectors vary greatly, and this variance must be closed. This variance has been recognised by participants in both sectors, and this is a positive step. NCURA has recognised this dimension through its program offerings.

Lastly, communication remains critical. This is a common sense conclusion, but if it was that easy, why hasn't communication been more effective? An analogy to the world of divorce law seems appropriate. As an attorney who has handled divorce cases, I can testify to the importance of tending to the entire relationship, not just discrete aspects of it. This is equally true of university-industry collaboration. It is my conclusion that the UIDP forum and TN will play integral roles in the continued strengthening of communication among and between university and industry partners.

Perhaps the following is also needed, as articulated by Thomas A. Stewart, editor of the *Harvard Business Review*:

You cannot manage for the long term unless you can make room in your head, and your company's collective head, to think, plan, and execute over a multiyear time span, even while tending to inevitable (and important) distractions.¹⁶

Does your university's leadership have these attributes? Does your company's leadership possess such attributes?

As the title of this chapter indicates, the UIDP and TN are incremental improvements to improving university-industry collaboration. It is hoped that this incremental process continues well into the future, to such an extent that it becomes second nature to develop and 'close the deal' on such partnerships.

¹⁶ Thomas A Stewart, 'What the Long Term Takes' (July-August 2007) *Harvard Business Review* 12.

CHAPTER FOURTEEN

STREAMLINING COLLABORATIVE AGREEMENTS IN AN e-RESEARCH WORLD

Anthony Austin and Professor Brian Fitzgerald¹

INTRODUCTION

On 22 January 2008, the Australian Minister for Innovation, Science and Research announced a review of the 'National Innovation System'² which intends to analyse the Australian innovation system and to 'build innovation capacity by bringing sectors, institutions and individuals together'.³

To achieve innovation through this style of collaboration, the different actors will inevitably need to engage with technologically enhanced research methods and practices known broadly as e-Research. The rapidly emerging e-Research landscape promises to accelerate the discovery of knowledge, to increase the access and dissemination of data and to provide the opportunity for the international and serendipitous exchange of knowledge. ⁴ The law will play a central role in this

¹ Professor Brian Fitzgerald is the Project Leader of the Legal Framework for e-Research Project (<<http://www.e-research.law.qut.edu.au/>>) and Anthony Austin is a Research Officer with the same project. We acknowledge the assistance of Steven Gething Research Officer and the contribution of Dr. Amanda McBratney who undertook research on e-research collaboration for the Legal Framework for e-Research Project in the first half of 2007 in helping us map out this landscape.

² See the Review of the National Innovation website at <<http://www.innovation.gov.au/innovationreview/Pages/home.aspx>>.

³ Senator the Honourable Minister Kim Carr speech to the Australian / Melbourne Institute 2008 Economic and Social Outlook Conference Economics and Commerce Building University of Melbourne 28 March. See <<http://minister.innovation.gov.au/SenatortheHonKimCarr/Pages/NEWAGENDAFORPROSPERITY.aspx>>.

⁴ The Legal Frameworks for e-Research Project Report, *Legal and project agreement issues in collaboration and e-Research: Survey Results* (2007) authored by Marce Heffernan and Nikki David

environment. It acts like an infrastructure to shape the flow of knowledge. In many collaborative projects, the negotiation and completion of agreements which outline the project are not only critical, but also represent one of the biggest barriers to effective collaboration. The purpose of this chapter is to consider how the negotiation and contractual frameworks for research can be streamlined to accommodate the coming era of collaborative e-Research.

STREAMLINING THE PROCESS

In a collaborative project, the law should be an enabler to innovation, not an inhibitor:

It is important that institutional arrangements are made so as to minimize the extent to which the law becomes an impediment to cooperation among researchers, whether directly or indirectly by undermining informal mechanisms of trust and dispute resolution.⁵

For the law to be an enabler, it must be supported by policies, principles and frameworks:

Perhaps the biggest problem facing e-Research is the lack of understanding and agreement as to what is required in terms of local and national information infrastructure to support e-Research activities. Without this common framework of understanding it is actually very difficult to come to legal agreement as to collaborative arrangements, sharing, and interaction beyond a narrow set of participants. This then actually inhibits the establishment of an open e-Research environment that starts to utilise the potential offered by digital technologies.⁶

and assisted by Dr Amanda McBratney, Scott Kiel Chisholm, Professor Brian Fitzgerald, Professor Anne Fitzgerald and Dr. John Abbot, 72. A copy of this report can be sourced at: <<http://eprints.qut.edu.au/archive/00009112/01/9112.pdf>>.

⁵ Professor Paul A. David, and Dr. Michael Spence, *Towards institutional infrastructures for e-Science: the scope of the challenge*, 7 at <<http://www.oii.ox.ac.uk/research/publications.cfm>>.

⁶ The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 62. See <<http://eprints.qut.edu.au/archive/00009112/01/9112.pdf>>.

The key question that arises is how to design and streamline the legal agreement process so that collaborative e-Research projects can be established and can commence without unnecessary delay?

To answer this question, four issues need to be considered:

- Institutional Frameworks and Policies. The establishment of institutional frameworks (potentially within existing government agencies) which should have input from industry, individuals and other organisations. This framework should facilitate the creation of national, state and localised policies that will promote the flow of knowledge that is necessary for collaboration to occur;
- Relational Frameworks. The creation of frameworks which clarify purposes and expectations between parties about collaborative projects and which engender trust and formulate relationships that are effective for collaboration to succeed;
- Tools. The design and employment of practical tools which have the effect of shortening the timeframe for the negotiation and drafting of formal collaboration agreements;
- Application. The implementation of institutional frameworks which provide training and education in these policies, relational frameworks and tools and that manages their adoption and utilisation by universities, industry and research.

This chapter addresses these four issues by examining:

- The perceptions of stakeholders in relation to collaborative projects and the processes for formalising collaborative project agreements;
- National and international initiatives and studies on proposed policies, frameworks and tools for facilitating collaborative agreements;
- The discussions held at the recent Queensland University of Technology Legal Framework for e-Research Roundtable; and
- Proposals for the streamlining of legal agreements for collaborative projects through institutional and relational frameworks, policies and tools.

STAKEHOLDER PERCEPTIONS – THE LEGAL FRAMEWORK FOR e-RESEARCH SURVEY – 2007

In order to assess the effectiveness of collaborative agreements, frameworks, tools and policies, it is necessary to understand how these issues are perceived by those who work in the collaborative e-Research environment.

In 2007, the Legal Frameworks for e-Research Project conducted a survey entitled *Legal and project agreement issues in collaboration and e-Research: Survey Results*.⁷ The survey obtained evidence from Australian researchers, research managers and legal advisors from universities, industry and government about legal and other issues in collaboration and e-Research.⁸ The survey sought to:

... identify common legal and project agreement problems encountered in forming research collaborations in order to from strategies to facilitate and streamline the process of e-Research in the Australian context.⁹

The survey concentrated on three specific themes:

Firstly, what are the legal procedures and norms for formalising collaborative e-Research agreements¹⁰ and how do these procedures and norms affect the parties and the success of collaborative innovation projects?¹¹

⁷ See *Legal and project agreement issues in collaboration and e-Research: Survey Results* at <<http://eprints.qut.edu.au/archive/00009112/01/9112.pdf>>.

⁸ 'e-Research' has been defined as: "...research activities that use a spectrum of advanced ICT capabilities and embraces new methodologies emerging from increased access to; Broadband communications networks, research instruments and facilities, sensor networks and data repositories; Software and infrastructure services that enable secure connectivity and interoperability; and Application tools that encompass discipline-specific tools and interaction tools...". The e-Research Coordinating Committee, *An Australian e-Research Strategy and Implementation Framework: Final report of the e-Research Coordinating Committee*, April 2006. See: <<http://www.dest.gov.au/NR/rdonlyres/3AC7BB72-3397-4269-A5FC-6758CDDFEF24/16579/eResearchFinalReportPublicVersionforweb.rtf>>

⁹ The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 8.

¹⁰ The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 25–31 and 37–42.

¹¹ The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 25–31 and 37–42.

Secondly, what are the problems encountered in negotiating issues of IP ownership, licensing, data access and what are other complications and delays that arise from formalising these agreements?¹² How do negotiations, complications and delays subsequently undermine feelings of trust and endanger the willingness of parties to participate in collaborative innovation?¹³

Thirdly, what are the participant's views on practical tools, relational frameworks and other strategies for simplifying the agreement process for collaborative e-Research projects?¹⁴

Survey participants were from research and management roles with most of them working in the university sector.¹⁵ A sizeable number of participants were involved in e-Research,¹⁶ stating that they are 'often' or are 'sometimes' are involved with different parties in collaborative research.¹⁷

¹² The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 43–46.

¹³ The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 46–54.

¹⁴ The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 55–62.

¹⁵ Of the 176 participants, 85 (or 48%) were in research roles, 66 (or 38%) were in research and/or organisational management and 25 (or 14%) were in legal or contracts roles. The majority of participants were from the University sector (64.8%), with 9.1% from Industry/Commercial and 9.1% from Government sectors, 10.8% from other Research Institutes and 6.3% from law firms. The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 14.

¹⁶ Approximately one-third (34.3%) of participants stated that they are 'extensively involved' with e-Research (37.1% moderately involved; 18.3% 'slightly involved' and 10.3% 'not at all involved'). Thirty-one percent of researchers, 41% of research/organisational managers, and 28% of the legal/contracts respondents stated that they are 'extensively involved' in e-Research. See Figure 3. Respondent's Involvement in e-Research, The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 15.

¹⁷ Universities 96%. Research institutions 85%. Industry participants 78%. See Figure 5. Relative Frequency of Involvement with Differing Parties, The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 21.

Support for Collaboration

Overall, the survey participants supported the concept of collaboration and in particular identified the attainment and the sharing of knowledge as being a major benefit of collaborative research. Participants ranked the importance of collaborative outcomes in the following order; the co-authoring of publications, the inflow of knowledge, the sharing of knowledge by public disclosure or publications, the improvement of research practices, the production of IP such as patents and copyright and obtaining access to improved work practices and better equipment or facilities.¹⁸ Benefits such as royalties, revenue, return on investment, licenses and start-up companies were less relevant outcomes for the participants.¹⁹ These results may reflect the academic nature of many survey participants.

Formal and Informal Collaboration

Many survey participants provided evidence of a strong culture of informal collaborative agreements and informal collaborative networks.²⁰ Less than half of the participants were involved in collaborations where formal collaborative agreements were entered into (such as master research agreements and licences).²¹

¹⁸ See Figure 8. Importance of Research Outcomes to Collaborative Projects, The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 32.

¹⁹ See Figure 8. Importance of Research Outcomes to Collaborative Projects, The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 32.

²⁰ 'Informal networks (including informal conversations, conference interactions)', 'informal agreements leading to co-authored publications' and 'single research contracts' were the most frequent arrangements cited. Approximately 70% of respondents stated that their collaborations often involve informal networks (including informal conversations, conference interactions). Only 7% of the sample stated that their collaborations often involve joint ventures, cross-licensing or technical assistance agreements. See Figure 7, Relative Frequency of Various Types of Collaboration Agreements/Arrangements, The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 26.

²¹ 44%. See Figure 7, Relative Frequency of Various Types of Collaboration Agreements/Arrangements, The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 26.

This preference for informal collaborative arrangements may stem from evidence that formal collaborative research agreements can take anything from three months for a simple two-party agreement to eight months for large, complex or multi-party agreements to be finalised.²²

Participants made particular comment about this issue:

“Legal agreements represent the largest impediment to timely research ...”²³

“... Unfortunately the formal agreements we use are becoming increasingly impractical due to the time and costs of developing the agreements ...”²⁴

“We had a 12-month ARC grant for which it took 15 months to get an MOU signed”²⁵

“Legal advice often tends to make the collaboration so formal /complicated that it endangers the willingness of collaborators to participate. Sometimes legal advice is too oriented towards protecting the interests of my organisation, so that it does not see that formal agreements need to be balanced win-win arrangements”²⁶

Survey participants also stated that given the timeframes of the parties, collaborative projects often commence before a formal collaborative agreement has been finalised:²⁷

²² The Legal Framework for e-Research Project’s Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 38.

²³ The Legal Framework for e-Research Project’s Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 38.

²⁴ The Legal Framework for e-Research Project’s Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 38.

²⁵ The Legal Framework for e-Research Project’s Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 51.

²⁶ The Legal Framework for e-Research Project’s Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 39.

²⁷ Commencing collaborative research projects prior to the signing of agreements is a relatively common practice; with 26% stating that they ‘often’ and 54.2% stating that they ‘sometimes’ commence projects before agreements are signed (only 6.8% stated that they never start projects prior to sign-off). The Legal Framework for e-Research Project’s Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 40.

“... you've got a short-ish timeline, and you can't afford to wait months for the haggling to stop. If you don't start before the contract is signed, you'll won't finish on time and end up in violation of the terms of agreement”²⁸

“The legal and contractual processes can often be much slower than the time it actually takes to complete the research!”²⁹

Many participants felt that there were certain issues that caused problems in the negotiation and the performance of formal collaborative research agreements which included; intellectual property-ownership; data ownership and access; intellectual property-licensing and the over-valuing of intellectual property.³⁰ These negotiation difficulties are perceived as eroding the feelings of trust between the participants:³¹

It is the mutual rapport and trust between parties that is vitally important. If there is no trust then even a perfectly good legal document may be misused ...³²

Tellingly, a majority of participants stated that the negotiation of a formal agreement ‘became too complex’³³ because of ‘differing expectations’³⁴ between the project parties and believed that negotiating with university technology transfer offices, industry and government

²⁸ The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 40 to 41.

²⁹ The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 41.

³⁰ See Figure 11. Specific Problems in Negotiating Formal Agreements, The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 47.

³¹ Over one-third of the sample (36%) stated that sometimes negotiation difficulties prevented the project from proceeding and that trust had been eroded. The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 43.

³² The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 55.

³³ See Figure 10. General Problems in Negotiating Formal Agreements, The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 43.

³⁴ See Figure 10. General Problems in Negotiating Formal Agreements, The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 43.

agencies carry with them their own particular difficulties when entering into formal agreements.³⁵

Despite the existence of a culture of informal collaboration, a sizeable number of survey participants indicated that they still want formal agreements for collaborative projects.³⁶ Interestingly, many participants stated that their collaborations never involve the need for patent protection or licensing arrangements.³⁷

Participants stated that for collaborative projects to succeed they needed shared goals, good relationships and communication with their project partners.³⁸ Most importantly, they wanted formal agreements that were easy to enter into,³⁹ particularly agreements which specifically addressed intellectual property, data ownership or data access and which exhibited a degree of flexibility in their application to collaborative projects.⁴⁰

³⁵ See Figure 10. General Problems in Negotiating Formal Agreements, The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 43.

³⁶ Almost one-third of the sample believe that formal agreements are always necessary (31.1%), with approximately two-thirds stating that formal agreements are sometimes necessary (68.0%). Over half of the sample (56.5%) also stated that they never conclude formal agreements without consultation or assistance. The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 39.

³⁷ Approximately one-in-three participants stated that their collaborations never involve patents, software, know-how or other intellectual property licences or Cooperative Research Centres. The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 25.

³⁸ Approximately half (49.0%) of comments made predominantly reflect the importance of research synergies and shared goals and resources, with approximately 40% of comments referring to the importance of good relationships and communication. The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 37.

³⁹ 92% of participants believed that formal agreements which were easy to enter into was 'somewhat' to 'very important' in order to increase collaborative e-Research. See Figure 14. Future Importance of Various Contracting Issues, The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 62.

⁴⁰ Intellectual property (53% stating that it will be 'very important' and 38% stating that it will be 'somewhat important'), 'Data ownership or access' (51% stating that it will be 'very important' and 42% stating that it will be 'somewhat important') and 'Flexibility of formal agreements' (43% stating that it will be 'very important' and 48% stating that it will be 'somewhat important'). See Figure 14. Future Importance of Various Contracting Issues, The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 62.

NATIONAL AND INTERNATIONAL INITIATIVES – PREVIOUS STUDIES

The issues raised by the survey report are reflective of concepts that have also been examined by significant Australian and overseas studies in the area of collaborative research. These studies themselves have many themes in common with each other, such as:

Links or partnerships between industry, universities and research institutions are necessary for increasing collaborative research,⁴¹ however issues regarding IP ownership and access are often viewed as impediments to collaboration;

There is a need for uniform and national approaches to IP ownership and licensing and establishing a set of best practice principles for industry and university collaboration with publicly funded research agencies;⁴² and

The final value of an output should be shared equitably, based on the direct proportional value of the inputs to a project.⁴³

Some of the recommendations, documents, guidelines or interactive tools which these studies have proposed are examined below. These proposals fall within four broad categories being;

- Technology enabled collaborative research agreements;
- Template collaborative research agreements;

⁴¹ The Prime Ministers Science, Engineering and Innovation Council (PMSEIC) Report *Australia's Science and Technology Priorities for Global Engagement*, December 2006, 61.

See

<http://www.dest.gov.au/sectors/science_innovation/publications_resources/profiles/Presentation_Global_Engagement.htm>.

⁴² Recommendation 11, Dr J Howard 2005, *Knowledge Exchange Networks in Australia's Innovation System: Overview and Strategic Analysis*, Howard Partners Pty Ltd, commissioned for The Business, Industry and Higher Education Collaboration Council (BIHECC). See <<http://www.dest.gov.au/NR/rdonlyres/D60AE27E-1BF3-4305-ACCC-3027FE0A43FF/8488/KENReportFinal.rtf>>.

⁴³ Department of Education, Science and Training, *Review of Closer Collaboration Between Universities and Major Publicly Funded Research Agencies*, 2004, 37. See <<http://www.dest.gov.au/NR/rdonlyres/327F4C1D-99CC-4F93-91FB-1A2DEA8F299E/3623/pub.pdf>>.

- Guidelines which advise parties on how to construct and draft collaborative research agreements for university - industry collaborations or to meet funding requirements;⁴⁴ and
- The creation of institutional frameworks which co-ordinate and facilitate the streamlining of legal processes for formal collaborative agreements.⁴⁵

TECHNOLOGY ENABLED COLLABORATIVE RESEARCH AGREEMENTS

The UIDP TurboNegotiator Project – 2006

A current project which is attracting much interest is the TurboNegotiator (“TN”), established by the University-Industry Demonstration Partnership (“UIDP”⁴⁶). The TN Project started in July 2006 and seeks to create an online methodology for constructing effective and equitable university-industry collaborative research agreements from clauses selected by the TN program in accordance with its *Guiding Principles for University-Industry Endeavours*.⁴⁷ These principles

⁴⁴ This has been adopted by the European Commission Seventh Framework Programme (See <http://cordis.europa.eu/fp7/home_en.html>), the Commission of the European Communities (See Commission of the European Communities, *Commission Staff Working Document Voluntary Guidelines for universities and other research institutions to improve their links with industry across Europe*, 2. (See <http://ec.europa.eu/invest-in-research/pdf/sec2007449_en.pdf>) and the CREST OMC Expert Group (See <http://ec.europa.eu/invest-in-research/policy/rd_collab_en.htm>).

⁴⁵ This is exemplified by Professor Paul A. David, and Dr. Michael Spence, *Towards institutional infrastructures for e-Science: the scope of the challenge*. See <<http://www.oii.ox.ac.uk/research/publications.cfm>>.

⁴⁶ The UIDP was established on 1 August 2006 has participating members from both industry and universities and is convened by The National Academies, Washington. UIDP developed out of the University-Industry Partnership Project (Established in 2003 and funded by the US National Council of University Research Administrators) Mayo, Merrilea J., *Current Status of University-Industry Relationships in the U.S. Innovation System*.

See <http://www.uidp.org/UIDP_Intro.pdf>, 2–3.

The University–Industry Congress of the National Council of University Research Administrators analysed examples of negotiations and collaborative projects between university and industry from August 2003 to April 2006. NCRA Report, *Living Studies in University-Industry Negotiations*, April 2006.

See <http://www7.nationalacademies.org/guirr/Guiding_Principles.pdf>.

⁴⁷ NCRA Report, *Guiding Principles for University-Industry Endeavours*, April 2006.

state that universities, research organisations and industry must understand:

- The various levels of their respective contributions to collaborative projects;⁴⁸
- Each other's objectives or 'missions', such as university interest in knowledge sharing and education and industry's interest in profitability;⁴⁹
- Their respective constraints in a collaborative project, particularly the need of industry partners for timely agreements which ensure appropriate commercial returns;⁵⁰ and
- The need to develop long-term relationships in collaborative research.⁵¹

TN was commenced following evidence showing that the negotiation of university-industry research agreements in the US is a barrier to research collaboration.⁵²

The TN online program is designed to balance each party's interests, contributions and constraints regarding a collaborative project. Agreements are formed which are tailored to the parties' interests, instead of them having to conform their interests to the parameters of an established template agreement.

Importantly, TN contains a 'project space' in which university and industry parties can obtain general consensus about each others perceptions and ideas for a collaborative project, before they select appropriate clauses for the research agreement. The project space includes 'consensus statements' which guide parties in dealing with contentious issues. Each statement comments on the issue at hand, explains the reasoning behind the statement, sets out 'principles' which

⁴⁸ Guiding Principle #1, 5–6. NCRA Report, *Guiding Principles for University-Industry Endeavours*, April 2006.

⁴⁹ Guiding Principle #1, 7. NCRA Report, *Guiding Principles for University-Industry Endeavours*, April 2006.

⁵⁰ Guiding Principle #1, 7. NCRA Report, *Guiding Principles for University-Industry Endeavours*, April 2006.

⁵¹ Guiding Principle #2. NCRA Report, *Guiding Principles for University-Industry Endeavours*, April 2006, 8.

⁵² Evidence presented at the University-Industry Congress Summit, Washington, 25 April 2006.

the parties must adhere to and lists other factors or comments that should be considered.⁵³

The TN software interview tool asks a series of questions of each party to the project. The questions are organised into specific sections relating to budgetary and investment considerations, the nature of the research, background IP, the probability of inventions resulting from the project, disclosure requirements, export controls, indemnification and potential conflicts of interest.⁵⁴ The answer provided to each question then determines the suggested agreement clauses. This tool will provide more than one alternative clause for the parties to choose from.⁵⁵

Clauses which are suggested then hyperlink to further information about their suitability in relation to each party's interests and their overall effect on the agreement and the project. Examples of such questions include: What is the nature of the project?; What are the disciplinary areas which the project encompasses?; What is the nature and extent of each parties contributions to the project?; What is the likelihood of a patentable result arising out of the project?; What are the costs to each party of participating and each party's market presence?; Who developed the research project concept and who made creative contributions to that concept?; Why do the parties want to work together?; Who owns background IP that could have impact on the project?; Who is funding the project?; What are the financial and non-financial contributions from the parties?; What are the types of labour and non-labour contributions

⁵³ There are draft consensus statements for how to produce a 'statement of work' of aims and activities for the project, how to determine issues of indemnification in the project, how to balance issues regarding the publication of project IP, dealing with copyrightable and other research results and dealing with background IP. See the drafts from the UIDP meeting April 9 to 10, 2008 at <http://www.osp.gatech.edu/TN/documents/ConsensusStmnt04_09_08.doc> and see presentation by Julia Garton, *TurboNegotiator, Milestones and Pathways, 23 July 2007*, Third Meeting of the University-Industry (Demonstration) Partnership, July 23–24, 2007. The National Academies Washington. See <http://www.uidp.org/UIDP_ARCHIVED_MEETINGS.html>.

⁵⁴ See the trial TurbNegotiator software (Limited to questions on IP issues) at: <<http://www.osp.gatech.edu/TN/index.html>>.

⁵⁵ Presentation by Julia Garton, *TurboNegotiator, Milestones and Pathways, 23 July 2007*, Third Meeting of the University-Industry (Demonstration) Partnership, July 23–24, 2007. The National Academies Washington. See <http://www.uidp.org/UIDP_ARCHIVED_MEETINGS.html>.

from the parties?; How important are they?; and Is there a need for confidentiality about the project?⁵⁶

TN will also measure how much time an agreement will take to finalise using the TN program. This data will be used to compare against how much time it takes to finalise non-TN facilitated agreements. It is still in the process of being developed and the UIDP has gone through several stages of development to date, incorporating member surveys, consultations and clinical software trials of the questionnaire program. The UIDP also hopes to develop a negotiation guide and/or a manual which will train negotiators in accordance with the program methodology and eventually, a national database of TN sourced data which can analyse negotiation trends and factors that impede negotiations. A working prototype of TN is expected to be available for internal UIDP trialling by December 2008.⁵⁷

TEMPLATE COLLABORATIVE RESEARCH AGREEMENTS

B-HERT – 1996

In contrast to the approach taken by UIDP, there have been a number of studies that advocate the use of template agreements for collaborative research projects.

In 1996, B-HERT⁵⁸ published a report, *Partners in Intellectual Property*,⁵⁹ which comparatively analysed the IP policies of certain higher education and research institutes. The report found that the interests of industry and the interests of universities and research differ in the development and commercialisation of IP. The objectives of universities and research

⁵⁶ See Casey, James J. Jnr, *The University-Industry Demonstration Partnership: An Incremental Improvement to University-Industry Collaboration*, Paper Presented at The Legal Framework for e-Research Conference, 11 to 12 July, 2007, Gold Coast, Australia.

⁵⁷ See the UIDP website at <http://www.uidp.org/UIDP_PROJECT_STATUS.html>.

⁵⁸ B-HERT is an association of Australian universities, corporations, professional associations and major public research organisations that seeks to "...pursue initiatives that will advance the goals and improve the performance of both business and higher education for the benefit of Australian society" by "...[influencing] public opinion and government policy on selected issues of importance". See <http://www.bhert.com/aboutBHERT_Mission.htm>.

⁵⁹ See *Partners in Intellectual Property* at <http://www.bhert.com/publications_Reports.htm>.

are largely identified as the advancement of knowledge as a contribution to society, whereas the objectives of industry are commercial and based upon achieving specific returns on investments.

The report also identified ‘friction points’ between universities, research and industry during negotiations about IP, including:

- Users’ rights and the reservation of rights to use the IP,⁶⁰ IP ownership issues regarding the ineffectiveness of joint ownership,⁶¹ royalty payments for improvements⁶² and competing interests between the use of exclusive licensing and assignments;⁶³
- Management issues such as profit sharing,⁶⁴ maintenance of communication⁶⁵ and relationships,⁶⁶ the rights of students to royalty income,⁶⁷ the status of project workers as inventors,⁶⁸ the publication of commercially sensitive work⁶⁹ and expectations for performance timeframes which are held by both parties;⁷⁰
- Cultural differences between industry and universities or research such as differences in negotiating and management styles,⁷¹ over reliance on legal expertise⁷² and the lack of experienced joint project supervisors;⁷³ and
- Unsuitable and inflexible contracts such as the over use of standard form contracts by industry or government

⁶⁰ *Partners in Intellectual Property*, 4.

⁶¹ *Partners in Intellectual Property*, 5.

⁶² *Partners in Intellectual Property*, 5.

⁶³ *Partners in Intellectual Property*, 6.

⁶⁴ *Partners in Intellectual Property*, 6.

⁶⁵ *Partners in Intellectual Property*, 6.

⁶⁶ *Partners in Intellectual Property*, 7.

⁶⁷ *Partners in Intellectual Property*, 7.

⁶⁸ *Partners in Intellectual Property*, 7.

⁶⁹ *Partners in Intellectual Property*, 7.

⁷⁰ *Partners in Intellectual Property*, 9.

⁷¹ *Partners in Intellectual Property*, 10.

⁷² *Partners in Intellectual Property*, 10.

⁷³ *Partners in Intellectual Property*, 10.

organisations in their dealings with universities and research institutions.⁷⁴

Importantly, the report proposes that collaboration parties enter into a 'Partnering Concept'⁷⁵ for the development and commercialisation of IP as a starting point for negotiations. The Partnering Concept discusses issues such as users' rights, IP ownership, maintaining communication and relationships and timing expectations.⁷⁶ It envisages three types of collaboration agreement scenarios:

- Universities or research organisations are the source of the background IP that is brought to the project;
- Industry is the source of the background IP that is brought to the project; and
- Where the background IP is still in the conceptual phase.⁷⁷

The report provides a contractual template for either the assignment or the licensing of IP rights called the 'Grant of Intellectual Property Rights'. It contains clauses that address IP licensing and assignment, consideration and warranties,⁷⁸ but it does not address the publication of IP results, management issues and contractual flexibility.

The Lambert Review - 2003

The use of template agreements for collaborative research projects was taken to a greater level of practical application by the U.K. Lambert Review in 2003.

In 2002, the United Kingdom Department for Education and Skills and the Department for Trade and Industry commissioned Richard Lambert to undertake a nationwide review of university and industry collaboration in the United Kingdom. The U.K. government was concerned that domestic business funded research was falling behind

⁷⁴ *Partners in Intellectual Property*, 11.

⁷⁵ *Partners in Intellectual Property*, 26.

⁷⁶ See clauses 3, 4, 7 and 8 and Annexure C to *Partners in Intellectual Property*.

⁷⁷ *Partners in Intellectual Property*, 27. It is unclear whether the Partnering Concept is intended to be a voluntary protocol or a binding agreement. Further, the report does not address issues of cultural differences or contractual flexibility.

⁷⁸ *Partners in Intellectual Property*, 40–48.

that of other major economies.⁷⁹ Consequentially, they wanted strategies to increase domestic and international business demand for British research and development and in particular, to improve the level of collaboration between industry and U.K. universities.⁸⁰

After its establishment in 2002, the Lambert Review of Business-University Collaboration went on to examine various barriers to increased industry and university collaboration, how they could be removed and how opportunities for collaboration could be increased. It identified case studies for industry-university collaborative ventures and analysed numerous stakeholder submissions on issues of collaborative research and government policy.

The Review issued *The Lambert Review of Business-University Collaboration - Final Report* in 2003. It contains thirty-three recommendations covering a broad range of policy strategies for facilitating knowledge transfer, third stream funding, university codes of governance and formal and informal networks between business people and academics. The Report specifically examined the role of collaborative research in promoting the transfer of knowledge between universities, industry and the wider community⁸¹ and in doing so identified:

... that collaborative research is one of the most effective forms of knowledge transfer.⁸²

It concluded that disagreements often arise in negotiations between industry and universities over the ownership of IP and exploitation rights, which were identified as time-consuming and expensive.⁸³ Failure to agree on IP ownership often deterred both industry and universities from research collaboration.⁸⁴ This is compounded where the parties fail to understand each other's intentions for the resulting IP, particularly where there is:

⁷⁹ See *The Lambert Review of Business-University Collaboration – Final Report* at <http://www.hm-treasury.gov.uk./media/9/0/lambert_review_final_450.pdf>, Chapter 2.

⁸⁰ *The Lambert Review of Business-University Collaboration – Final Report*, 9–10.

⁸¹ *The Lambert Review of Business-University Collaboration – Final Report*, Chapter 3.

⁸² *The Lambert Review of Business-University Collaboration – Final Report*, paragraph 3.31.

⁸³ *The Lambert Review of Business-University Collaboration – Final Report*, 3.34–3.36 and 4.12 to 4.18.

⁸⁴ *The Lambert Review of Business-University Collaboration – Final Report*, 3.34–3.36 and 4.12 to 4.18.

... no clear framework ... to help the two sides balance their competing interests.⁸⁵

The report firstly recommended that a set of model collaborative research agreements be created and used on a voluntary basis in order to speed up negotiations in university-industry collaborative projects.⁸⁶

Secondly, an 'IP protocol' should be established between industry and universities as a starting point for negotiation. Under the protocol, universities would automatically own the IP arising from collaborative research and industry would be able to negotiate the licensing of this IP. Industry could still own this IP whenever it makes significant contributions to the collaborative project.⁸⁷ Regardless of how IP ownership is determined, the protocol requires that universities must not be restricted in their future research capabilities, business must develop IP applications in a timely manner and the substantive results of the research must be published within an agreed period.⁸⁸

In 2004, the Lambert Working Group was established. It was chaired by Richard Lambert and included stakeholders from industry and university bodies. The Lambert Working Group developed five model research collaboration agreements (and supporting materials) known as 'Model Agreements'.⁸⁹

Model Agreements One, Two and Three are designed to start negotiations between university and industry from the position that:

- The university owns the IP that results from the project.⁹⁰ The university is free to publish about the IP or have its staff and students discuss the project in tutorials or lectures⁹¹(unless business or industry has issued a 'confidentiality notice' to

⁸⁵ *The Lambert Review of Business-University Collaboration – Final Report*, 3.34–3.36 and 4.12 to 4.18.

⁸⁶ *The Lambert Review of Business-University Collaboration – Final Report*, 3.37.

⁸⁷ *The Lambert Review of Business-University Collaboration – Final Report*, 4.19 - 4.27.

⁸⁸ *The Lambert Review of Business-University Collaboration – Final Report*, Recommendation 4.1.

⁸⁹ See the Lambert Model Agreements at

<<http://www.innovation.gov.uk/lambertagreements/index.asp?lv1=2&lv2=0&lv3=0&lv4=0>>.

⁹⁰ For example, clause 4.3. Model Collaborative Research Agreement No1.

⁹¹ For example, clause 5.1. Model Collaborative Research Agreement No1.

- prevent publication and discussion until patent or other protection for the IP has first been obtained);⁹²
- Each party retains ownership in their own IP which they bring to the project,⁹³ but grants to each other a royalty-free, non-exclusive licence to use this ‘background’ IP only for the purposes of the project;⁹⁴
 - The university grants a non-exclusive licence to industry to use the IP resulting from the project for any purpose within an agreed ‘field’⁹⁵ or territory (indefinite, fully paid-up and royalty free);⁹⁶
 - The non-exclusive licence allows industry to sub-licence the IP, provided it is to employees or agents and it is for the purposes of the project;⁹⁷
 - The information, techniques or know-how which each party brings to the project cannot be disclosed to third parties;⁹⁸
 - A university will not be in breach of confidence by publishing or permitting discussion of IP, provided that they have not first received a ‘confidentiality notice’ from industry.⁹⁹ This notice is designed to protect confidential information regarding business and industry and to minimise any risk to the possibility of obtaining a patent or other protection for the IP results.¹⁰⁰

⁹² For example, clause 5.2. Model Collaborative Research Agreement No1.

⁹³ For example, clause 4.1. Model Collaborative Research Agreement No1.

⁹⁴ For example, clause 4.2. Model Collaborative Research Agreement No1.

⁹⁵ Meaning a specific business or technological area. See the definition of ‘The Field and the Territory’ in the Lambert Agreements Guidance Notes at

<<http://www.innovation.gov.uk/lambertagreements/index.asp?lv11=3&lv12=0&lv13=0&lv14=0#note9>>.

⁹⁶ For example, clause 4.5. Model Collaborative Research Agreement No1.

⁹⁷ For example, clause 4.5. Model Collaborative Research Agreement No1.

⁹⁸ For example, clause 6. Model Collaborative Research Agreement No1.

⁹⁹ For example, clause 6.3. Model Collaborative Research Agreement No1.

¹⁰⁰ See the sample confidentiality notice at

<http://www.innovation.gov.uk/lambertagreements/files/Sample_Confidentiality_Notice.DO>.

Model Agreement One is the basic non-exclusive licence model. Models Two and Three repeat the provisions of Model Agreement One and only differ from it in further providing: an option for industry to negotiate an exclusive license rights for IP¹⁰¹ or an option for industry to take an assignment of IP.¹⁰² In both Model Agreements Two and Three, the university still retains the right to use the IP for academic teaching and research.¹⁰³

Model Agreements 4 and 5 are designed to start negotiations between university and industry from the position that:

- Industry owns the IP resulting from the project.¹⁰⁴ The university or any student or contractor must assign any rights they have in the resulting IP to industry;¹⁰⁵
- Each party retains ownership in their own IP which they bring to the project,¹⁰⁶ but grants to each other a royalty-free, non-exclusive licence to use each others 'background' IP only for the purposes of the project;¹⁰⁷
- Unlike Model Agreements 1 to 3, industry does not grant universities a non-exclusive licence to use the resulting IP outside of the actual project.¹⁰⁸

¹⁰¹ For example, clause 4.6. Model Collaborative Research Agreement No.2.

See

<http://www.innovation.gov.uk/lambertagreements/files/Lambert_Agreement_2_Ink.doc>.

¹⁰² For example, clause 4.6. Model Collaborative Research Agreement No.3.

See

<http://www.innovation.gov.uk/lambertagreements/files/Lambert_Agreement_3_Ink.doc>.

¹⁰³ For example, clause 4.7. Model Collaborative Research Agreements No.2 and No.3.

¹⁰⁴ For example, clause 4.3. Model Collaborative Research Agreements No.4 and No.5.

See

<http://www.innovation.gov.uk/lambertagreements/files/Lambert_Agreement_4_Ink.doc>

and

<http://www.innovation.gov.uk/lambertagreements/files/Lambert_Agreement_5_Ink.doc>.

¹⁰⁵ For example, clauses 4.3 and 4.4. Model Collaborative Research Agreements No.4 and 5.

¹⁰⁶ For example, clause 4.1. Model Collaborative Research Agreements No.4 and 5.

¹⁰⁷ For example, clause 4.2. Model Collaborative Research Agreements No.4 and 5.

¹⁰⁸ For example, clause 4.6. Model Collaborative Research Agreements No.4 and 5.

Model Agreement 4 provides the ability for universities to still publish and disseminate the IP for the ‘advancement of education through teaching and research’¹⁰⁹ (subject to the terms of any ‘confidentiality notice’ issued by industry).¹¹⁰

Under Model Agreement 5, the university has no publication or dissemination rights as in Model Agreement 4¹¹¹ and can only use resulting IP for the purposes of the project itself.¹¹²

The Lambert Working Group has also supplied an ‘Outline’, consisting of questions designed to prompt the parties to think about and to discuss with each-other certain issues about the project before they select one of the model agreements, being; financial contributions, background IP, the project results, confidentiality and publication, liability and termination.¹¹³

A ‘Decision Guide’¹¹⁴ is also available for use in connection with the agreements. The guide provides a series of questions designed to determine which of the five agreements is best suited for the project at hand, based on each party’s answers to those questions. The questions focus on issues such as; reliance on background IP by the parties and the need for access to background IP;¹¹⁵ the need for universities to publish results and the need of sponsors to countenance publication;¹¹⁶ which

¹⁰⁹ For example, clause 5.1. Model Collaborative Research Agreements No.4.

¹¹⁰ For example, clause 5.2. Model Collaborative Research Agreements No.4.

¹¹¹ See the Outline of the Lambert Agreements at

<<http://www.innovation.gov.uk/lambertagreements/index.asp?lv1=4&lv2=0&lv3=0&lv4=0>>.

¹¹² For example, clause 4.6. Model Collaborative Research Agreement 5.

¹¹³ See the Outline of the Lambert Agreements at

<<http://www.innovation.gov.uk/lambertagreements/index.asp?lv1=4&lv2=0&lv3=0&lv4=0>>.

¹¹⁴ See the Outline of the Lambert Agreements at

<<http://www.innovation.gov.uk/lambertagreements/index.asp?lv1=2&lv2=1&lv3=0&lv4=0>>.

¹¹⁵ Sections 1, 2 and 4. The Lambert Agreements Decision Guide.

See

<<http://www.innovation.gov.uk/lambertagreements/index.asp?lv1=2&lv2=1&lv3=0&lv4=0>>.

¹¹⁶ Section 2. The Lambert Agreements Decision Guide.

parties have lead the projects, the relevancy of results to each party; the likelihood of patentable results and the likelihood of results that industry may not be interested in;¹¹⁷ the need for exclusive licences; funding and budget considerations;¹¹⁸ who was the catalyst for the project; what the parties' interest in the project is and what are the parties financial and non-financial contributions to the project.¹¹⁹

'Guidance Notes' are also available which provide plain English definitions of the defined terms used in the agreements and explanations about the effect and intention of certain clauses.¹²⁰

CRC INC - 2006

The concept of template collaborative research agreements was also considered by the Australian Institute for Commercialisation ("AIC"¹²¹) and the Cooperative Research Centres Association ("CRCA"¹²²) who in 2006 produced a 'Model Constitution Document' and a 'Model Participants Agreement' for use where a CRC is being formed as a joint venture company limited by guarantee.

The AIC and the CRCA recommend that:

... the template documentation should be treated as a starting point and each CRC and its participants must seek their own professional legal, accounting and taxation advice to determine whether they appropriately address the objectives and risks applicable to their own CRC.¹²³

¹¹⁷ Section 3. The Lambert Agreements Decision Guide.

¹¹⁸ Section 4. The Lambert Agreements Decision Guide.

¹¹⁹ Additional Questions. The Lambert Agreements Decision Guide.

¹²⁰ The Lambert Agreements Decision Guide.

¹²¹ The AIC is a private organisation that provides consultancy services in the technology transfer facilitation and brokerage of intellectual property. See the AIC website at <http://www.ausicom.com/01_cms/details.asp?ID=19>.

¹²² The Cooperative Research Centres Association is the umbrella organisation for the 56 Cooperative Research Centres ("CRCs") that operate in Australia in six industry areas. The stated aim of the CRC Programme (administered by DEEWR) is to "...enhance Australia's industrial, commercial and economic growth through the development of sustained, user-driven, cooperative public-private research centres that achieve high levels of outcomes in adoption and commercialisation". See <http://www.crca.asn.au/about_crcs/default.htm>.

¹²³ See the AIC website at <http://www.ausicom.com/01_cms/details.asp?ID=624>.

The ‘Constitution Document’ is a company constitution document and details standard procedures for the CRC company’s organisation, including membership, general meetings, voting, the appointment and removal of directors and the powers and remuneration of directors.

The ‘Participants Agreement’ is a template contractual agreement between the ‘Participants’¹²⁴ and the CRC company. The intellectual property clauses set out rights and obligations about the use of background IP, the ownership of CRC IP and its commercialisation. All background IP is licensed by participants to each other and to the CRC company and depends on whether it will be used for either the project, for commercialisation or for general use. IP generated by a CRC project can be owned in accordance with the following options:¹²⁵

- The CRC company owns the project IP; or
- The Participants and the CRC company will own the beneficial interest in the project IP as tenants in common in accordance with the ‘Project Shares’¹²⁶ or in equal shares if no ‘Project Shares’ are specified; or
- The CRC company owns the interest of the Participants in the project IP on trust.

This agreement is primarily designed for the commercialisation of resulting project IP by the CRC company because it has an exclusive right to commercialise the project IP and grant licences.¹²⁷ Non-company Participants must obtain a licence to use project IP,¹²⁸ unless otherwise authorised.¹²⁹ They must provide information about project IP to the company when requested and must not deal with CRC IP in any way unless authorised.¹³⁰ Non-company Participants must grant the company a perpetual, irrevocable, royalty-free, non-exclusive licence for

¹²⁴ Participants are those persons or bodies (other than the Company) who sign the Participants Agreement. See the definition of “Participants”. Clause 1.1 of the Participants Agreement.

¹²⁵ Clause 22.1 of the Participants Agreement.

¹²⁶ Being the proportional entitlement of Participants and the CRC company as set out in the agreement. See the definition of “Project Shares” Clause 1.1 of the Participants Agreement.

¹²⁷ Clauses 23.1 and 23.2. Participants Agreement.

¹²⁸ Ibid, Clause 22.6. Participants Agreement.

¹²⁹ Ibid, Clause 22.15. Participants Agreement.

¹³⁰ Ibid, Clause 22.13. Participants Agreement.

any improvements which non-company Participants make to the project IP.¹³¹

This agreement also requires that a ‘Commercialisation Plan’ must be circulated to all participants before the project IP is exploited.¹³² Non-company Participants cannot commercialise project IP and can only use it for teaching purposes or for internal research, provided that this use does not impede upon designated confidential information or the ability to protect and commercialise resulting IP.¹³³

GUIDELINES

CREST - 2006

In addition to the issue of how collaborative research agreements are to be created, a number of studies have developed guidelines or toolkits which will assist parties in choosing and constructing these agreements.

This issue was examined by the CREST OMC 2nd Cycle Expert Group on Intellectual Property¹³⁴ in 2006 when it published their report *Cross-Border Collaboration between Publicly Funded Research Organisations and Industry and Technology Transfer Training*.¹³⁵ CREST sought to produce guidelines which improve the ‘coherence and effectiveness’ of IP ownership rights that are ‘applicable in publicly funded research’¹³⁶ and to develop

¹³¹ Ibid, Clause 22.16. Participants Agreement.

¹³² Ibid, Clause 23.4. Participants Agreement.

¹³³ Ibid, Clause 22.15. Participants Agreement.

¹³⁴ This group was established in 2005 and is one of the five expert groups created by CREST, the European Union’s Scientific and Technical Research Committee. This group consists of members from various European government departments, patent offices and the European Commission.

¹³⁵ See *Report of the CREST OMC Expert Group on Intellectual Property 92nd Cycle*, *Cross-Border Collaboration between Publicly Funded Research Organisations and Industry and Technology Transfer Training* at <http://ec.europa.eu/invest-in-research/pdf/download_en/crestreport.pdf>.

¹³⁶ *Report of the CREST OMC Expert Group on Intellectual Property 92nd Cycle*, *Cross-Border Collaboration between Publicly Funded Research Organisations and Industry and Technology Transfer Training*, 9.

methodologies for improving and facilitating cross-border collaborative projects.¹³⁷

The report proposed that a toolkit be adopted to enable a collaborating party to identify how IP issues are handled in another European Union member state. The toolkit is designed to assist parties to make a decision about the best strategy for determining ownership of and access to the IP resulting from a project. It does this by providing explanation and commentary on ownership and rights to use IP, financial contributions made by industry, the university's use of results of academic purposes and cross-border differences and legal requirements for other project partner's jurisdictions.

The CREST toolkit is currently active,¹³⁸ although it is still under development and is subject to a review at the end of 2008. It consists of a 'First Step' which is an interactive checklist of questions that users answer. The questions relate to deciding a suitable position for ownership of the IP rights and provide answers based on a proportion of 'Yes' responses on a scale of one to ten. The questions look at issues such as; the importance of results for future activity; exploitation of the results; who conceived the project?, what is the purpose of the project? and why the respective industry and university parties want to fund or carry out the project?¹³⁹ It also provides a 'Second Step' for guidance on cross-border issues regarding IP rights and ownership, negotiations, funding, confidentiality, publication and the protection and enforcement of IP rights.¹⁴⁰ This 'Second Step' is meant to be used in conjunction with 'Fact Sheets' in relation to each member state. The 'Fact Sheets' explain:

¹³⁷ *Report of the CREST OMC Expert Group on Intellectual Property 92nd Cycle), Cross-Border Collaboration between Publicly Funded Research Organisations and Industry and Technology Transfer Training*, 9.

¹³⁸ See the CREST Interactive Toolkit at <http://ec.europa.eu/invest-in-research/policy/rd_collab_en.htm>.

¹³⁹ See the First Step. CREST Interactive Toolkit at <http://ec.europa.eu/invest-in-research/policy/rd_collab_en.htm>.

¹⁴⁰ See the Second Step. CREST Interactive Toolkit at <http://ec.europa.eu/invest-in-research/pdf/download_en/secnd_step.pdf>.

- Types of IP rights which can be obtained from universities/research institutes¹⁴¹ in another member state¹⁴² and their particular rules or requirements regarding confidentiality or publication;
- Who owns the IP rights at these universities/research institutes, the legal situation regarding IP rights derived from public funding and differences between the member states that impact on the ownership of IP rights;
- Who is entitled to negotiate IP contracts at universities/research institutes, what are the terms on which IP rights can be obtained and at what price;
- How funding affects IP ownership and exploitation and any relevant tax effects that impact on funding;
- Specific requirements regarding IP rights, who will pay for the costs of obtaining them, who will enforce them and links to further information about IP rights.¹⁴³

The toolkit contains a 'Decision Guide' which proposes five ownership positions that the project parties could adopt, being:

- The university owns the IP and grants a non-exclusive licence to industry to use the IP in a specific field or geographical area;
- As above, with industry having a right to negotiate to acquire an exclusive licence to certain IP;
- As above, with industry having a right to negotiate to take ownership of the IP through an assignment;
- Industry owns the IP with university reserving a right to use IP for teaching, research and publication, subject to confidentiality conditions; or
- As above, but the university has no right to publish the IP.¹⁴⁴

¹⁴¹ *Report of the CREST OMC Expert Group on Intellectual Property 92nd Cycle) Cross-Border Collaboration Between Publicly Funded Research Organisations and Industry and Technology Transfer Training*, 16.

¹⁴² Such as patents copyrights, trademarks or designs.

¹⁴³ See the appendices to the *Report of the CREST OMC Expert Group on Intellectual Property 92nd Cycle) Cross-Border Collaboration Between Publicly Funded Research Organisations and Industry and Technology Transfer Training*, 101.

Interestingly, these ownership positions have degrees of similarity to the ownership positions under the Lambert Model Agreements.

The toolkit also includes the 'Intellectual Property Right Interactive Visualisation Tool'. This software tool enables users to select two member state countries and then obtain; a comparison of legislative and legal positions between the two countries in relation to the types of IP rights available in each country; each states position on the ownership of IP rights and the negotiation of IP rights contracts; the effect of funding on IP rights contracts; confidentiality and publication; and the protection and enforcement of IP rights. This particular tool also links to country specific websites about government activities and national laws which are relevant to these issues.¹⁴⁵

Commission of the European Communities Voluntary Guidelines – 2007

Collaborative guidelines were also set down by the Commission of the European Communities who in 2007 produced a Commission Staff Working Document as a response to a survey into cooperation and knowledge transfer between universities, research institutes and industry.¹⁴⁶

The Commission recommended guidelines for developing a standard approach for the management and transfer of knowledge and intellectual property regarding publicly funded collaboration projects.¹⁴⁷ These guidelines have established 'good practices' for publicly funded

¹⁴⁴ The CREST Cross-Border Collaboration Decision Guide, 7. See <http://ec.europa.eu/invest-in-research/pdf/download_en/introd.pdf>.

¹⁴⁵ See the CREST country comparative questionnaire programme at <<http://ec.europa.eu/invest-in-research/policy/tool.htm>>.

¹⁴⁶ The European Commission, Directorate-General for Research 2006 online survey, *Draft Report on the Outcomes of the Public Consultation On Transnational Research Cooperation And Knowledge Transfer Between Public Research Organisations And Industry* ('EC Knowledge Transfer Report'), 1 September 2006, at <http://ec.europa.eu/invest-in-research/pdf/download_en/consult_report.pdf>.

¹⁴⁷ See *Commission of the European Communities, Commission Staff Working Document Voluntary Guidelines for Universities and Other Research Institutions to Improve their Links with Industry Across Europe*, 2 at <http://ec.europa.eu/invest-in-research/pdf/sec2007449_en.pdf>.

collaborative research contracts in Europe. Some of the practices emphasise personal relationships, openness and compatibility with the universities goals or 'mission' and recommends use of the CREST decision guide. They also focus on: communication in negotiations to avoid misunderstandings; a clear delineation of rights between the parties; the ownership of IP and access rights and determining the likely commercial applications of the project from the outset; identification of financial and non-financial input to a project by the respective parties; clear discussion regarding the nature and scope of a project, the protection for IP rights; the impact on each others future research; the usage of model contracts and whether model contracts will permit negotiation on background IP, ownership, confidentiality issues and access rights.¹⁴⁸

The guidelines also contain general advice on non-exclusive licensing or assignments and advocates that universities and research institutions should reserve the right to publish IP results in collaborative agreements. It advises that they should only keep IP results confidential, subject to 'detailed assessment and justification'.¹⁴⁹

FP7 – 2007:

The European Commission Seventh Framework Programme ("FP7") is an interesting example of a study which at first tried to create uniform agreements solely from guidelines, without providing a draft agreement for reference purposes. However, a group of FP7 stakeholders later created a draft template agreement to assist parties to comply with those guidelines.

FP7 commenced in 2007¹⁵⁰ and sought to make the European Union a dynamic competitive knowledge-based economy¹⁵¹ through a combination of 'research, education and innovation'.¹⁵² FP7 provides

¹⁴⁸ *Commission of the European Communities, Commission Staff Working Document Voluntary Guidelines for Universities and Other Research Institutions to Improve their Links with Industry Across Europe*, 10–12.

¹⁴⁹ *Commission of the European Communities, Commission Staff Working Document Voluntary Guidelines for Universities and Other Research Institutions to Improve their Links with Industry Across Europe*, 14–15.

¹⁵⁰ The European Union implements numerous 'framework programmes' to support research activities in the European Union of which FP7 is the latest such programme.

¹⁵¹ See the FP7 website at <http://ec.europa.eu/research/leaflets/fp7/page_03_en.html>.

¹⁵² See the FP7 website at <http://ec.europa.eu/research/leaflets/fp7/page_03_en.html>.

funding grants, in accordance with established regulations, to collaborative projects involving researchers, research centres, universities and other entities.¹⁵³ The regulations are mandatory upon parties who wish to obtain FP7 funding and it provides guidelines to drafting a collaborative research agreement.¹⁵⁴ The guidelines and regulations include the following:

- Resulting IP is owned by those participants who generated it. Where respective shares are unable to be ascertained, the parties shall have joint ownership,¹⁵⁵ unless they agree on a different solution. Resulting IP must be protected by the owner¹⁵⁶ or else the European Commission may take ownership.¹⁵⁷ Resulting IP can be transferred,¹⁵⁸ however the Commission can prevent transfer if it is not in accordance with developing the competitiveness of the European economy.¹⁵⁹ Commercial use will only be undertaken for valid commercial reasons;¹⁶⁰
- FP7 funding recipients must use and disseminate the resulting IP,¹⁶¹ providing that the parties have made a decision about possible IP protection¹⁶² and confidentiality.¹⁶³ Interestingly, there is no express prohibition in the regulations or the guidelines against publication of the resulting IP;

¹⁵³ See Regulation (EC) No 1906/2006 of the European Parliament and of the Council of 18 December 2006 laying down rules for the participation of undertakings, research centres and universities in actions under the Seventh Framework Programme and for the dissemination of research results (2007–2013), Official Journal of the European Union L 391/1, 30.12.2006 at <<http://cordis.europa.eu/documents/documentlibrary/90798681EN6.pdf>>.

¹⁵⁴ See the Guide to Intellectual Property Rules for FP7 Projects Version 28/06/2007 at <ftp://ftp.cordis.europa.eu/pub/fp7/docs/ipr_en.pdf>.

¹⁵⁵ Article 39. Guide to Intellectual Property Rules for FP7 Projects Version 28/06/2007.

¹⁵⁶ Article 44. Guide to Intellectual Property Rules for FP7 Projects Version 28/06/2007.

¹⁵⁷ Guide to Intellectual Property Rules for FP7 Projects Version 28/06/2007, 10.

¹⁵⁸ Article 42. Guide to Intellectual Property Rules for FP7 Projects Version 28/06/2007.

¹⁵⁹ Guide to Intellectual Property Rules for FP7 Projects Version 28/06/2007, 11.

¹⁶⁰ Guide to Intellectual Property Rules for FP7 Projects Version 28/06/2007, 13.

¹⁶¹ Guide to Intellectual Property Rules for FP7 Projects Version 28/06/2007, 23.

¹⁶² Guide to Intellectual Property Rules for FP7 Projects Version 28/06/2007, 14.

¹⁶³ Article 46. Guide to Intellectual Property Rules for FP7 Projects Version 28/06/2007.

- Exclusive licences can be granted for both resulting and background IP¹⁶⁴ and the parties must have access to any parties background IP if that is necessary for them to enjoy use of the resulting IP;¹⁶⁵ and
- Licences and third party user rights can be granted if they are necessary for the project¹⁶⁶ and can be granted on a royalty free basis.¹⁶⁷ However, the Commission can reverse licences to third parties if they are deemed detrimental to European competitive advantage.¹⁶⁸

In order to obtain the benefit of FP7 funding for a collaborative project,¹⁶⁹ most participants must enter into and adhere to a 'FP7 Model Grant Agreement'¹⁷⁰ and a 'FP7 Model Consortium Agreement'.¹⁷¹ The 'Model Grant Agreement' sets out the terms of funding.¹⁷² The 'Negotiation Guidance Notes' explains how participants should apply for and negotiate with the Commission for a 'Grant Agreement'.¹⁷³ Responsibility for drafting the 'Consortium Agreement' lies with the project parties and they must do so in accordance with the parameters of the regulations and the requirements for FP7 funding under the 'Checklist for a Consortium Agreement for FP7 Projects'¹⁷⁴ and the 'Guide to Intellectual Property Rules for FP7 Projects'.¹⁷⁵

¹⁶⁴ Article 48. Guide to Intellectual Property Rules for FP7 Projects Version 28/06/2007.

¹⁶⁵ Article 50. Guide to Intellectual Property Rules for FP7 Projects Version 28/06/2007.

¹⁶⁶ Guide to Intellectual Property Rules for FP7 Projects Version 28/06/2007, 14.

¹⁶⁷ Guide to Intellectual Property Rules for FP7 Projects Version 28/06/2007, 49.

¹⁶⁸ Guide to Intellectual Property Rules for FP7 Projects Version 28/06/2007, 22.

¹⁶⁹ Article 1. Guide to Intellectual Property Rules for FP7 Projects Version 28/06/2007.

¹⁷⁰ Articles 18–19. Guide to Intellectual Property Rules for FP7 Projects Version 28/06/2007.

¹⁷¹ Article 24. Guide to Intellectual Property Rules for FP7 Projects Version 28/06/2007.

¹⁷² See the Model Grant Agreement at <http://cordis.europa.eu/fp7/calls-grant-agreement_en.html#standard_ga>.

¹⁷³ See the Negotiation Guidance Notes at <ftp://ftp.cordis.europa.eu/pub/fp7/docs/negotiation_en.pdf>.

¹⁷⁴ See the Checklist for a consortium Agreement for FP7 Projects at <ftp://ftp.cordis.europa.eu/pub/fp7/docs/checklist_en.pdf>.

¹⁷⁵ See the Guide to Intellectual Property Rules for FP7 Projects at <ftp://ftp.cordis.europa.eu/pub/fp7/docs/ipr_en.pdf>.

However, despite the availability of guidelines, the DESCAs group (which was initiated by FP7 stakeholders¹⁷⁶) has subsequently produced a simplified consortium agreement which is intended to balance all interests of all partners in an FP7 project.¹⁷⁷ It is supplied as a draft template only and DESCAs emphasises that the FP7 regulations still need to be taken into account by project parties. The template is set out in a comparative table format with the suggested clause in the left hand column and in the right hand column, an explanation of definitions and the effect of the clause. It often provides optional clauses for parties to choose with explanations as to the effect of each option¹⁷⁸. The options reflect the preferences of stakeholder research organisations and universities; however the template warns that mixing the options can cause inconsistencies in the agreement.¹⁷⁹

DESCAs further provides four illustrative examples of the template for use in situations concerning; a small project on fair and reasonable conditions;¹⁸⁰ a small project based on royalty free access;¹⁸¹ a large project based on fair and reasonable conditions;¹⁸² and a large project based on royalty free access.¹⁸³

¹⁷⁶ See the DESCAs Core Group website at <<http://www.desca-fp7.eu/the-desca-core-group/>>.

¹⁷⁷ See the DESCAs Simplified FP7 Model Consortium Agreement Version 2 at <http://www.desca-fp7.eu/fileadmin/content/Documents/DESCA__version__2_final.doc>.

¹⁷⁸ Clause 8.1. DESCAs Simplified FP7 Model Consortium Agreement Version 2.

¹⁷⁹ See the DESCAs Core Group website <<http://www.desca-fp7.eu/the-desca-consultation/>>, Remark 4.

¹⁸⁰ See “small project” : “fair and reasonable conditions” at <http://www.desca-fp7.eu/fileadmin/content/Documents/DESCA__version__2_SP_O1.doc>.

¹⁸¹ See “small project” : “royalty free access” at <http://www.desca-fp7.eu/fileadmin/content/Documents/DESCA__version__1_May_2007_example2_GOV_SP_OP2.doc>.

¹⁸² See “large project” : “fair and reasonable conditions” at <http://www.desca-fp7.eu/fileadmin/content/Documents/DESCA__version__1_May_2007_example3_GOV_LP_OP1.doc>.

¹⁸³ See “large project” : “royalty-free access” at <http://www.desca-fp7.eu/fileadmin/content/Documents/DESCA__version__1_May_2007_example4_GOV_LP_OP2.doc>.

CREATION OF INSTITUTIONAL FRAMEWORKS

McGauchie – 2004

Finally, certain studies have examined the need for over-arching institutional frameworks which co-ordinate and facilitate the utilisation of agreements, guidelines and tools and which also develop national policies on collaborative endeavours.

In March 2004, the Australian Federal Government Department of Education, Science and Training (DEST, now DEEWR) published a report entitled *Review of Closer Collaboration between Universities and Major Publicly Funded Research Agencies* (“The McGauchie Review”).¹⁸⁴ The report contained the findings of a review committee (chaired by Donald McGauchie) who convened to examine the potential to exploit collaboration between Publicly Funded Research Agencies (“PFRAs”) and universities and possible models for closer collaboration.¹⁸⁵

The report’s review committee defined collaboration as a ‘partnership, alliance or network aimed at a mutually beneficial clearly defined outcome’¹⁸⁶ and it describes various benefits¹⁸⁷ and barriers¹⁸⁸ to collaboration, its drivers and models¹⁸⁹ and how collaboration could be enhanced¹⁹⁰ through co-location, networking and clustering. The committee identified what they believed were key barriers to collaboration between business and universities or PFRA bodies, including:¹⁹¹

- Cultural differences between PFRA and universities – with a respective ‘industry- pull’ and ‘research-push’;¹⁹²

¹⁸⁴ Department of Education, Science and Training, *Review of Closer Collaboration Between Universities and Major Publicly Funded Research Agencies*, 2004.

See <http://www.dest.gov.au/NR/rdonlyres/42A4E965-16F1-4614-965E-11966D66D8EC/3624/issues_paper.pdf>.

¹⁸⁵ *Review of Closer Collaboration Between Universities and Major Publicly Funded Research Agencies*, 39.

¹⁸⁶ *Review of Closer Collaboration Between Universities and Major Publicly Funded Research Agencies*, 1.

¹⁸⁷ *Review of Closer Collaboration Between Universities and Major Publicly Funded Research Agencies*, 2.

¹⁸⁸ *Review of Closer Collaboration Between Universities and Major Publicly Funded Research Agencies*, 2.

¹⁸⁹ *Review of Closer Collaboration Between Universities and Major Publicly Funded Research Agencies*, 4.

¹⁹⁰ *Review of Closer Collaboration Between Universities and Major Publicly Funded Research Agencies*, 5.

¹⁹¹ *Review of Closer Collaboration Between Universities and Major Publicly Funded Research Agencies*, 33.

¹⁹² *Review of Closer Collaboration Between Universities and Major Publicly Funded Research Agencies*, 33.

- Limited access to finance, the level of entrepreneurial management skills available, the predominance of a risk adverse culture and the availability of business and finance expertise;
- A lack of significant tax incentives for businesses who invest in university research and the need for specific funding if commercialisation is a requirement; and
- IP issues¹⁹³ including IP ownership, contractual disputes, overvaluation of IP, the need for a clear set of principles or policies for IP management¹⁹⁴ and the cost of IP protection.

The committee concluded that some institutions spend ‘significant energy on detailed up-front negotiation of [the] IP issue’.¹⁹⁵ Protracted negotiations over IP ownership and exploitation were unnecessary at the outset of a collaborative programme¹⁹⁶ and the committee found that it is only in a small number of instances that research output reaches the stage for commercialisation.¹⁹⁷

The report suggests that parties should defer detailed negotiations on IP exploitation issues until specific milestones in the research have been reached, commercialisation prospects improve or the collaboration has matured so the contributions of each party can be more accurately determined.¹⁹⁸

In order to resolve protracted negotiations over IP ownership, the report states that the IP resulting from a collaborative project will need to be shared between the parties based on their proportional contribution to the project:¹⁹⁹

“Intellectual property, generated as a result of collaborative research, should be divided according to the relative inputs of the various collaborators. The inputs must be measured by their demonstrable relevance to the generated property.

¹⁹³ *Review of Closer Collaboration Between Universities and Major Publicly Funded Research Agencies*, 36.

¹⁹⁴ *Review of Closer Collaboration Between Universities and Major Publicly Funded Research Agencies*, 36.

¹⁹⁵ See also the UIDP’s comments on triaging “agreements into high/low probability of generating valuable IP”: *Living Studies in University-Industry Negotiations*, April 2006, 16.

¹⁹⁶ *Review of Closer Collaboration Between Universities and Major Publicly Funded Research Agencies*, 36.

¹⁹⁷ *Review of Closer Collaboration Between Universities and Major Publicly Funded Research Agencies*, 36.

¹⁹⁸ *Review of Closer Collaboration Between Universities and Major Publicly Funded Research Agencies*, 36.

¹⁹⁹ *Review of Closer Collaboration Between Universities and Major Publicly Funded Research Agencies*, 36.

Consideration should be given to better utilisation of existing commercial arbitration and mediation mechanisms to handle and resulting disputes”.²⁰⁰

Importantly, the report also recommended that the Federal Government establish a ‘Strategic Research Council’ to provide them with policy advice about collaboration and which will implement a set of ‘National Research Priorities’;²⁰¹ a ‘Framework’ to measure the performance of publicly funded research agencies and universities in order to encourage collaboration;²⁰² a ‘Collaboration Fund’ to finance collaborative projects between business and industry and universities and research institutions;²⁰³ and a clear set of national principles or policy for IP management.²⁰⁴

BIHECC – 2005

The McGauchie Review was followed on by a BIHECC²⁰⁵ commissioned report in 2005 to investigate ‘knowledge exchange networks’, described as:

... structured intermediary mechanisms for users to locate, exchange and acquire knowledge in a systematic way, with a view to development of new products, processes and services.²⁰⁶

²⁰⁰ *Review of Closer Collaboration Between Universities and Major Publicly Funded Research Agencies*, 37.

²⁰¹ *Review of Closer Collaboration Between Universities and Major Publicly Funded Research Agencies*, 15.

²⁰² *Review of Closer Collaboration Between Universities and Major Publicly Funded Research Agencies*, 23.

²⁰³ *Review of Closer Collaboration Between Universities and Major Publicly Funded Research Agencies*, 26.

²⁰⁴ Which they believed were not being met by the 2001 *National Principles of IP Property Management for Publicly Funded Research*. *Review of Closer Collaboration Between Universities and Major Publicly Funded Research Agencies*, page xi.

²⁰⁵ The Business, Industry and Higher Education Collaboration Council (BIHECC) was established in 2004 and provides advice to the Australian Federal Government Minister for Education, Employment and Workplace Relations on ways to increase collaboration between the higher education sector and other public and private business, industry, community and educational organisations.

²⁰⁶ Dr. J. Howard 2005, *Knowledge Exchange Networks in Australia's Innovation System: Overview and Strategic Analysis*, Howard Partners Pty Ltd, at

<http://www.dest.gov.au/sectors/science_innovation/publications_resources/profiles/ken.htm>.

See the Australian Federal Government's Productivity Commission's Review of Public Support for Science and Innovation at

The report describes the various communication channels that exist between researchers, developers and industry and made thirteen recommendations to the Australian Federal Government's Productivity Commission's Review of Public Support for Science and Innovation,²⁰⁷ which included:

- Establishing a separate source of public funding to support knowledge transfer and pre-commercialisation activities of universities;²⁰⁸
- Better incentives for pre-commercialisation collaboration, the early stages of commercialisation and for exploiting publicly funded research and development in order to increase collaboration and knowledge transfer;²⁰⁹
- Supporting knowledge brokering infrastructure to link up institutions and industry and supporting knowledge exchange networks between industry, universities and research;²¹⁰ and
- Publicly funding science and innovation to provide longer term funding for research and knowledge transfer.²¹¹

The report did not call for an institutional framework to implement these recommendations, but some of the recommendations are ones which may need to be created and administered by such a framework, such as; uniform national approaches to IP ownership and licensing;²¹² public policies which balance IP protection for publicly funded research outcomes;²¹³ and establishing a set of best practice principles for collaboration between industry, universities and PFRAs.²¹⁴

<<http://www.dest.gov.au/NR/rdonlyres/E929FA3D-0F29-40E4-A53B-65715083C54D/8489/KENReportFinal.pdf>>.

²⁰⁷ Productivity Commission 2007, *Public Support for Science and Innovation*, Research Report, Productivity Commission, Canberra. See <http://www.pc.gov.au/__data/assets/pdf_file/0016/37123/science.pdf>.

²⁰⁸ Recommendation 1. *Public Support for Science and Innovation*.

²⁰⁹ Recommendation 2. *Public Support for Science and Innovation*.

²¹⁰ Recommendation 5. *Public Support for Science and Innovation*.

²¹¹ Recommendation 13. *Public Support for Science and Innovation*.

²¹² Recommendation 7. *Public Support for Science and Innovation*.

²¹³ Recommendation 7. *Public Support for Science and Innovation*.

²¹⁴ Recommendation 11. Recommendation 7. *Public Support for Science and Innovation*.

David and Spence – 2003

In what has become a landmark report in the area of e-Research, *Towards institutional infrastructures for e-Science: the scope of the challenge*²¹⁵ advocates the creation of institutional frameworks for streamlining the agreement process in collaborative projects.

This report released in 2003 examined the legal issues and processes associated with collaborative projects in the U.K. It identified that collaborative projects are often organised on an informal basis, rather than being defined by a written signed agreement²¹⁶ and that as a consequence may not be enforceable at law.²¹⁷ Furthermore, the report goes on to state that standard form contracts are not effective in providing what the parties want in a collaborative project or in making allowances for actual research practices or in establishing a degree of trust between project parties.²¹⁸

The report suggests that standard form contracts exacerbate the problems raised by informal collaborations, because such contracts rarely re-set the terms of the agreement each time the project circumstances change²¹⁹ and cannot include subsequent parties to the contract without having to re-make the entire agreement.²²⁰ Other problems associated with standard form contracts include:

- An increased legal risk for the parties because standard form contracts are often used without forethought as to their appropriateness to the project at hand or without reference to appropriate legal advice;²²¹
- Standard form contracts may impede the commencement of projects because the parties are unable to choose between each others standard form contracts;²²²

²¹⁵ By Paul A. David and Michael Spence. See *Towards institutional infrastructures for e-Science: the scope of the challenge* at <<http://www.oii.ox.ac.uk/research/publications.cfm>>.

²¹⁶ *Towards institutional infrastructures for e-Science: the scope of the challenge*, 38.

²¹⁷ *Towards institutional infrastructures for e-Science: the scope of the challenge*, 38.

²¹⁸ *Towards institutional infrastructures for e-Science: the scope of the challenge*, 8–9.

²¹⁹ *Towards institutional infrastructures for e-Science: the scope of the challenge*, 38.

²²⁰ *Towards institutional infrastructures for e-Science: the scope of the challenge*, 38.

²²¹ *Towards institutional infrastructures for e-Science: the scope of the challenge*, 55.

²²² *Towards institutional infrastructures for e-Science: the scope of the challenge*, 55.

- Allowing industry parties to exercise unequal bargaining power or pressure over and against the interests of university or research parties.²²³

Furthermore, standard form contracts set the parameters of a collaborative project before the project commences, which acts against establishing relationships of trust between the parties²²⁴ and may have the effect of stifling project research practices.²²⁵ Whilst legal advisors may have the ability to draft contracts on the behalf of parties, they are not in a position to comprehend all of the issues for all parties (only for the party which they represent) and cannot objectively balance and reconcile the contending interests and risks for both universities and industry.²²⁶

This report recommends that a public agency be established which will co-ordinate and facilitate solutions²²⁷ and whose main task will be to provide a menu of ready made agreement clauses. These clauses can be selected by parties to the project to resolve specific problems in their collaboration project agreement.²²⁸ Because parties select their own clauses, the agreement is built ‘from the ground up’, with clauses reflecting each parties true project intentions and avoids problems caused by standard template contracts.

The report proposes that a public agency (known as the ‘Advisory Board on Collaboration Agreements’) be established which will produce, evaluate and update standard contractual clauses. The clauses are intended to apply to various types of problems or situations that arise in collaborative research projects²²⁹ and will be able to be assembled into a

²²³ *Towards institutional infrastructures for e-Science: the scope of the challenge*, 55.

²²⁴ *Towards institutional infrastructures for e-Science: the scope of the challenge*, 10.

²²⁵ *Towards institutional infrastructures for e-Science: the scope of the challenge*, 54.

²²⁶ *Towards institutional infrastructures for e-Science: the scope of the challenge*, 30.

²²⁷ *Towards institutional infrastructures for e-Science: the scope of the challenge*, 52.

²²⁸ *Towards institutional infrastructures for e-Science: the scope of the challenge*, 28. Survey participants showed some agreement for the creation of a government agency to develop and maintain a master database of standard clauses for research contracts, issue guidelines and oversee licensing practices, oversee licensing practices. See Figure 12. Ways to Improve the Negotiation Process. The Legal Framework for e-Research Project’s Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 56.

²²⁹ *Towards institutional infrastructures for e-Science: the scope of the challenge*, 51.

variety of alternating collaboration agreements.²³⁰ The advisory board will also determine and draft a set of underlying principles for their functionality and applicability.²³¹

European Research Area Expert Group Knowledge Sharing Recommendations - 2008

The European Research Area Expert Group recently issued a report²³² which merges the concepts of policies, guidelines and model agreements in order to produce a European wide approach for a knowledge sharing system between publicly funded research organisations (“PROs”) and industry.²³³

The report recommends that the Commission of the European Communities²³⁴ issues a Recommendation to European Union member states to implement certain strategies in order to facilitate the sharing of knowledge generated by public funding and to ensure that industry and PRO collaboration agreements are ‘put into place more quickly and smoothly and to reduce transaction cost’.²³⁵ Whilst supportive of the Commission’s voluntary guidelines for PROs to improve their links with industry,²³⁶ the report recommends the creation of guidelines that are

²³⁰ *Towards institutional infrastructures for e-Science: the scope of the challenge*, 53.

²³¹ *Towards institutional infrastructures for e-Science: the scope of the challenge*, 41 and 53. To date, the U.K. government has not established such an agency.

²³² *The Report of the European Research Area Expert Group: Knowledge Sharing in the European Research Area*, 2008 at <http://ec.europa.eu/research/era/pdf/era-gp-cg4_en.pdf>. This is the Final Report of one of seven Expert Groups established by the DG Research of the European Commission as a follow-up to the Green Paper, *The European Research Area: New Perspectives*, published in April 2007.

²³³ *Report of the European Research Area Expert Group: Knowledge Sharing in the European Research Area*, 13.

²³⁴ See also *The Commission of the European Communities, Commission Staff Working Document Voluntary Guidelines for Universities and Other Research Institutions to Improve their Links with Industry Across Europe*.

²³⁵ *Report of the European Research Area Expert Group: Knowledge Sharing in the European Research Area*, 26.

²³⁶ See also *The Commission of the European Communities, Commission Staff Working Document Voluntary Guidelines for Universities and Other Research Institutions to Improve their Links with Industry Across Europe*.

aimed at and adhered to by both PROs and industry²³⁷ combined with relational frameworks where each party focuses on ‘realistic expectations of what might be achieved’²³⁸ in a collaboration. Other relevant strategies include:

- A voluntary ‘Knowledge Sharing Code of Practice’ and a model form of IP Policy,²³⁹ which will operate as a ‘reference point for those collaborating or intending to collaborate with European PROs and for European PROs collaborating or hoping to collaborate with industry’.²⁴⁰ This code and policy will ‘raise the awareness of European PROs of the need for them to engage in knowledge sharing and to manage knowledge effectively, to set out a set of minimum standards which European PROs may adopt on a voluntary basis and, by doing so, to facilitate interaction between European PROs and industry’;²⁴¹

²³⁷ “...the guidelines are targeted at PROs, but it takes two or more to form a contract. In the context of practices relating to the ownership of, and access to, intellectual property rights financial, human and intellectual input, the exploitation of intellectual property rights, confidentiality, the enforcement of intellectual property rights and relationship management it is essential that all parties (be they PROs or industry) abide by the same practices; PROs cannot implement these guidelines unless industry is also willing to implement them”. *Report of the European Research Area Expert Group: Knowledge Sharing in the European Research Area*, 13.

²³⁸ “Engaging in knowledge sharing is a contact sport and should not be a war. PROs need to appreciate that industry may have to put a lot of effort in before the results of the research are ready to be exploited and the intellectual property created by PROs may not have the immediate value the PRO supposes. Neither PROs nor industry should indulge in negotiations for the sake of winning every point, no matter how unimportant; both should have realistic expectations of what might be achieved”. *Report of the European Research Area Expert Group: Knowledge Sharing in the European Research Area*, 13.

²³⁹ *Report of the European Research Area Expert Group: Knowledge Sharing in the European Research Area*, 26.

²⁴⁰ *Report of the European Research Area Expert Group: Knowledge Sharing in the European Research Area*, 26.

²⁴¹ The proposed code must at least address the following issues: “A defined position of responsibility for overseeing knowledge sharing activities within the PRO; A clear position on the ownership of intellectual property rights created by PRO staff; Procedures for identifying and notifying intellectual property rights capable of commercial application; A mechanism for assessing the potential interest in intellectual property rights capable of commercial application, taking account of social, economic and enforcement conditions that prevail in the relevant territory and sector; The systematic use of records of the creation of intellectual property rights, such as laboratory notebooks; Mechanisms to deal with actual and potential conflicts of interest; A policy regarding publication of the results of research, taking into account situations when publication must be or should be delayed or withheld, and for how long, and, if

- The training of professional staff in technology transfer to ensure that collaborations happen within a shorter timeframe;²⁴²
- The harmonisation of funding conditions in relation to ownership and exploitation of intellectual property, so that PROs can negotiate appropriate terms for assignment or licensing of IP with industry;²⁴³ and
- The widespread adoption of model agreements and guidelines by PRO's and industry as exemplified by the Lambert Review agreements and toolkits.

LEGAL FRAMEWORK FOR e-RESEARCH ROUNDTABLE 12–13 JUNE 2008

As demonstrated from our examination of the survey and various Australian and overseas studies, many policies, frameworks and practical tools have been proposed to streamline the agreement process.

Despite some differences of approach, these proposals have a common goal; to produce a high degree of clarity between parties as to their respective contributions, duties and entitlements in a collaborative project.²⁴⁴ This common goal and the views, issues, frameworks, policies

applicable, clear delineation between the intellectual property rights owned by staff and those owned retained by the PRO; Clear lines of responsibility for procedure and policy management; Appropriate and clear timescales in respect of knowledge sharing and knowledge management procedures; A clear description of the rights and responsibilities of staff in relation to third party intellectual property rights; A clear description of the rights and responsibilities of students in relation to intellectual property rights created by students, and of staff in relation to those intellectual property rights; and a requirement to identify the PRO's contribution to knowledge wherever possible, for example within academic publications, and as the address for service for inventors employed by the PRO in any patent applications". *Report of the European Research Area Expert Group: Knowledge Sharing in the European Research Area*, 42 to 44.

²⁴² *Report of the European Research Area Expert Group: Knowledge Sharing in the European Research Area*, 26.

²⁴³ *Report of the European Research Area Expert Group: Knowledge Sharing in the European Research Area*, 27.

²⁴⁴ As certain survey participants commented "Problems often arise because the parties do not properly communicate and therefore they are not aware that they may have different expectations" and "Clarity between partners at the outset reduces the potential for later disagreement. The agreement need not be complex. Undue complexity is the major disincentive to developing formal agreements". The Legal Framework for e-Research Project's Report,

and tools raised and suggested by the survey report and the studies discussed previously were recently examined in a Roundtable held by the Queensland University of Technology Faculty of Law.²⁴⁵

Prior to the Roundtable, the Review of the National Innovation System received over 600 submissions from stakeholders regarding the future of innovation in Australia. Three of those submissions from the Legal Framework for e-Research Project,²⁴⁶ the Group of Eight Universities²⁴⁷ and the CSIRO²⁴⁸ addressed issues regarding the streamlining of collaborative research agreements. The Group of Eight submission specifically examined:

- Disputes over intellectual property ownership and licensing and rights in relation to background IP;
- The seeking of unreasonable warranties and indemnities;
- The right to disseminate research results in a timely manner;
- “No conflict of interest” clauses;
- The scope of suspension and termination rights;
- Respect for moral rights.

These issues were further discussed by the Roundtable.

Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results, 38 and 51.

²⁴⁵ The Legal Framework for e-Research Roundtable Workshop: Streamlining Collaboration in an e-Research World, held at the Queensland University of Technology 12–13 June 2008.

²⁴⁶ See the Legal Framework for e-Research Project’s submission to the National Innovation Review, *Streamlining Negotiation and Contracting in Collaborative Research Environments* at the National Innovation Review website at <http://www.innovation.gov.au/innovationreview/Documents/428A-Brian_Fitzgerald_and_Anthony_Austin.pdf>.

²⁴⁷ See the Group of Eight’s submission to the National Innovation Review, *In the Interests of Innovation: Time for a New Approach to Negotiating Research Agreements between the Commonwealth and Australian Universities* at the National Innovation Review website at <http://www.innovation.gov.au/innovationreview/Documents/372-Group_of_Eight_Supporting.pdf>.

²⁴⁸ See the CSIRO’s submission to the National Innovation Review, *Supplementary Submission from the CSIRO to the National Innovation System Review: Agreements Between the Australian Government and Publicly Funded Research Agencies* at <<http://www.innovation.gov.au/innovationreview/Documents/217A-CSIRO.pdf>>.

The proposals from the Roundtable suggested:

- A set of national policies on collaboration and knowledge sharing which underpins e-Research;
- A clear set of national principles or policies regarding IP ownership and licensing for collaborative projects;²⁴⁹
- Funding policies and conditions need a degree of uniformity across all funding agencies for collaborative projects, which would then create a national system of funding agreements.²⁵⁰ Uniform policies would also ensure that parties in negotiations with funding agencies will provide collaboration agreements that are in line with funding conditions;²⁵¹
- Parties to collaborative projects need to undertake realistic assessments of risk when negotiating collaboration agreements. Many collaboration agreements are often delayed because of protracted negotiations about issues that could otherwise be assumed as a reasonable risk, about issues such as background IP and warranties and indemnities;²⁵²
- Commensurate with the realistic assumption of risk, parties must be able to distinguish between vital objectives (whether the project will generate valuable IP or not) and irrelevancies²⁵³ and to devote time to complex collaborations instead of every several single transaction using up the resources of the

²⁴⁹ Reflective of the McGauchie Review. See Department of Education, Science and Training, *Review of Closer Collaboration Between Universities and Major Publicly Funded Research Agencies*, 2004, 36.

²⁵⁰ As discussed at The Legal Framework for e-Research Roundtable Workshop: Streamlining Collaboration in an e-Research World, held at the Queensland University of Technology 12–13 June 2008.

²⁵¹ As discussed at The Legal Framework for e-Research Roundtable Workshop: Streamlining Collaboration in an e-Research World, held at the Queensland University of Technology 12–13 June 2008.

²⁵² As discussed at The Legal Framework for e-Research Roundtable Workshop: Streamlining Collaboration in an e-Research World, held at the Queensland University of Technology 12–13 June 2008.

²⁵³ As discussed at The Legal Framework for e-Research Roundtable Workshop: Streamlining Collaboration in an e-Research World, held at the Queensland University of Technology 12–13 June 2008.

parties.²⁵⁴ This approach was supported by survey participants who advocated the ‘triaging’ of collaborative agreements for negotiation into those agreements that need significant negotiation and those which do not.²⁵⁵ Practically speaking, agreements of low risk and value to parties could be reduced to one to two page templates, instead of being made subject to detailed negotiation and review;²⁵⁶

- The need for a statement of national principles and guidelines to assist the implementation of a database of clauses and/or template agreements which will lead to the creation of a single national best practice resource;
- The utilisation of a ‘terms sheets’ for the agreement of key principles between the parties for specific types of transactions,²⁵⁷ which provide a plain English understanding of each party’s respective ideas, objectives, roles, commitments and expectations regarding a collaborative project, before the parties begin negotiations for an agreement;
- Trust must be established between parties through a ‘pre-agreement space’, where the parties are required to meet several times to discuss a potential collaboration, before they even begin to negotiate agreement terms;²⁵⁸

²⁵⁴ As discussed at The Legal Framework for e-Research Roundtable Workshop: Streamlining Collaboration in an e-Research World, held at the Queensland University of Technology 12–13 June 2008.

²⁵⁵ Figure 12. Ways to Improve the Negotiation Process. The Legal Framework for e-Research Project’s Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 56.

²⁵⁶ An example of this is the CSIRO’s FastTrack contracting system (<<http://www.csiro.au/org/ps9l.html>>) which focuses on simple non-disclosure agreements, testing agreements or postgraduate scholarships agreements. See the CSIRO example postgraduate scholarship agreement at <<http://www.csiro.au/files/files/p2za.pdf>>.

²⁵⁷ As discussed at The Legal Framework for e-Research Roundtable Workshop: Streamlining Collaboration in an e-Research World, held at the Queensland University of Technology 12–13 June 2008.

²⁵⁸ As discussed at The Legal Framework for e-Research Roundtable Workshop: Streamlining Collaboration in an e-Research World, held at the Queensland University of Technology 12–13 June 2008.

- Standard template agreements for use in collaborative projects²⁵⁹ that are intended to shorten negotiation timeframes and to remove delays caused by each party dissecting each other's standard agreements.²⁶⁰ This can be best exemplified by the Lambert Agreements or standard agreements which can be customised to the intentions and purposes of the collaborative parties.²⁶¹ However, it is important to remember that it is likely that there will never be a template agreement that will be designed to suit every collaborative situation. Template agreements can only be utilised as a starting point that saves a certain amount of negotiation time, not as the reduction of the agreement process to a software tool, and must be accompanied by guidance notes, decision guides or other similar material which forces the parties to address all issues required for a collaboration agreement;²⁶²
- The assembly and formation of agreements from a database of standard clauses. This was proposed by the UIDP TurboNegotiator project and was also viewed by survey participants as a practical tool for streamlining.²⁶³ If this idea was encapsulated in the form of a national database of standard

²⁵⁹ 68% of participants 'agreeing' or 'strongly agreeing'. See Figure 6. Ways to Streamline the Documentation Process. The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 59–61.

²⁶⁰ Also known as the 'battle of the forms'. As discussed at The Legal Framework for e-Research Roundtable Workshop: Streamlining Collaboration in an e-Research World, held at the Queensland University of Technology 12–13 June 2008.

²⁶¹ This was an option favoured by survey participants: 89% of participants 'agreeing' or 'strongly agreeing'. See Figure 6. Ways to Streamline the Documentation Process. The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 59–61. Survey participants also showed a preference for; template agreements which allowed the details on collaborative projects to simply be added on (87% of participants 'agreeing' or 'strongly agreeing'); licensing agreements based on the free open source software model (75% of participants 'agreeing' or 'strongly agreeing') and simple confidentiality agreements (86% of participants 'agreeing' or 'strongly agreeing').

²⁶² As exemplified by the Lambert Agreements and as discussed at The Legal Framework for e-Research Roundtable Workshop: Streamlining Collaboration in an e-Research World, held at the Queensland University of Technology 12–13 June 2008.

²⁶³ 76% of participants 'agreeing' or 'strongly agreeing'. See Figure 6. Ways to Streamline the Documentation Process. The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 59–61.

clauses, it may help to create an organic system of uniform agreements in the Australian collaboration environment;

- Practical tools and policies will be of limited use if researchers, research managers and other parties do not receive education and training about these tools, policies and the basics of IP and contractual law.²⁶⁴ The majority of survey participants were in favour of an increase in educational materials, guidelines and skills training for knowledge engagement practitioners.²⁶⁵ Commentators have advocated the creation of a specific tertiary level course in technology transfer, which would instruct how agreements for collaborative projects can be streamlined.²⁶⁶ Education and training in itself may also help to engender feelings of participation and vested interests in these tools and policies from industry, universities and research;²⁶⁷
- Collaborations are often frustrated because there is confusion within parties about their project objectives or because they do not identify who is authorised to negotiate with other interests.²⁶⁸ It is important that parties have established frameworks for their own internal communication and decision processes,²⁶⁹ have resolved any internal issues regarding IP

²⁶⁴ As discussed at The Legal Framework for e-Research Roundtable Workshop: Streamlining Collaboration in an e-Research World, held at the Queensland University of Technology 12–13 June 2008.

²⁶⁵ 71% of participants ‘agreeing’ or ‘strongly agreeing’. See Figure 12. Ways to Improve the Negotiation Process. The Legal Framework for e-Research Project’s Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 56.

²⁶⁶ As discussed at The Legal Framework for e-Research Roundtable Workshop: Streamlining Collaboration in an e-Research World, held at the Queensland University of Technology 12–13 June 2008.

²⁶⁷ As discussed at The Legal Framework for e-Research Roundtable Workshop: Streamlining Collaboration in an e-Research World, held at the Queensland University of Technology 12–13 June 2008.

²⁶⁸ As discussed at The Legal Framework for e-Research Roundtable Workshop: Streamlining Collaboration in an e-Research World, held at the Queensland University of Technology 12–13 June 2008.

²⁶⁹ As discussed at The Legal Framework for e-Research Roundtable Workshop: Streamlining Collaboration in an e-Research World, held at the Queensland University of Technology 12–13 June 2008.

ownership²⁷⁰ and have a clear intellectual property policy that balances issues of IP ownership, access, cost recovery and return on investment²⁷¹ before they enter into a collaborative agreement. The survey participants advocated a working rule that intellectual property generated in collaborative research should be divided according to relative inputs of the parties, measured by demonstrable relevance to the generated property;²⁷²

- The re-invigoration of existing institutional frameworks to train negotiators to balance and resolve issues from the position of a neutral adjudicator in proposed collaborative agreements. Such frameworks would co-ordinate the use of practical tools, policies and supporting materials at either a state or federal level or both. This would be supported by a national cross-sectoral legal advisory group that designs legal and policy frameworks and aligns appropriate methodologies for the streamlining of collaborative research agreements.

CONCLUSION

The survey results, the studies and the roundtable raise many questions and issues for consideration. As we have seen from the survey report, it can take up to 8 months to conclude a formal agreement because legal procedures and norms for formalising such agreements can delay and even stifle collaborative projects. The prolonged negotiation of agreement issues, such as the ownership and access rights for resulting intellectual property, reach through rights into each parties background IP and the extent of indemnities and warranties²⁷³ often leads to delays

²⁷⁰ As discussed at The Legal Framework for e-Research Roundtable Workshop: Streamlining Collaboration in an e-Research World, held at the Queensland University of Technology 12–13 June 2008.

²⁷¹ Figure 12. Ways to Improve the Negotiation Process. The Legal Framework for e-Research Project's Report, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results*, 56.

²⁷² As discussed at The Legal Framework for e-Research Roundtable Workshop: Streamlining Collaboration in an e-Research World, held at the Queensland University of Technology 12–13 June 2008.

²⁷³ As discussed at The Legal Framework for e-Research Roundtable Workshop: Streamlining Collaboration in an e-Research World, held at the Queensland University of Technology 12–13 June 2008.

and complications that undermine trust and the willingness of parties to collaborate.

It is hoped that the issues discussed in this chapter may be considered by the Australian Government as part of the Review of the National Innovation System and that they can add to the valuable work being done by technology transfer officers, research managers, researchers and legal advisors to streamline agreement processes for collaborative projects.

Universities, industry and researchers need to be able to shorten the timeframe for formalising collaborative research agreements. Parties want to collaborate on innovative projects at the time when their interest, motivation and utilisation of resources will be at its height.

Whilst e-Research is an excellent technology for collaborative projects, the technology alone will not enable collaboration to occur. If collaborative innovation is to prosper, then what is required in Australia is the synchronised and institutional development of policies, relational frameworks and practical tools for streamlining collaborative e-Research project agreements.²⁷⁴

²⁷⁴ See the Legal Framework for e-Research Project's submission to the National Innovation Review, *Streamlining Negotiation and Contracting in Collaborative Research Environments* at the National Innovation Review website at http://www.innovation.gov.au/innovationreview/Documents/428A-Brian_Fitzgerald_and_Anthony_Austin.pdf.

PART SIX

PRIVACY AND e-RESEARCH

A WIN:WIN FOR DATA ACCESS: BALANCING PUBLIC GOOD WITH PRIVACY CONCERNS

Professor Fiona Stanley AC¹

INTRODUCTION

This chapter suggests that the current urgent issues facing modern societies demand the best information and knowledge from which decisions can be made. This is vital for governments at all levels, non-government organisations and researchers whose work is used by those making decisions and policy. Such information is commonly available but rarely used, linked, re-used and analysed intelligently to inform such decision-making. As many problems are global, finding, sharing and analysing such data in robust national and international collaborations are essential activities. Such problems include environmental degradation, climate change, global pandemics, increases in obesity and mental ill health, overpopulation and city planning, water, security, crime and youth unrest.

A recent report entitled *From Data to Wisdom*, prepared for the Prime Minister's Science, Engineering and Innovation Council (PMSEIC) made several recommendations to put Australia in a strong position to both monitor and analyse these pressing problems internally and to be at the international table, when appropriate, to participate in planning and evaluating global threats.²

¹ FAA, FASSA, MSc, MD, FFPHM, FAFPHM, FRACP, FRANZCOG, Hon DSc, Hon DUniv, Hon FRACGP, Hon MD, Hon FRCPCH; Director, Telethon Institute for Child Health Research; Professor, School of Paediatrics and Child Health, The University of Western Australia.

² R Batterham et al, Prime Minister's Science, Engineering and Innovation Council, Working Group on Data for Science, *From data to wisdom: Pathways to successful data management for Australian Science* (2006) <www.dest.gov.au>.

One major issue in population data linkage is the balance between using individual health records on the total population for important public good activities, while at the same time ensuring that such private information is kept confidential. The rationale for using such data includes obtaining accurate and unbiased assessments of risks of disease and the effects of medical care. A win:win process to allow access and to protect privacy that has been developed and used in Western Australia for over 30 years is described below.

RATIONALE FOR DATA SHARING AND e-RESEARCH – THE PMSEIC REPORT

The science data challenges facing Australia and all other countries include:

1. The exponential increase in data assets and how they can be turned into knowledge and wisdom;
2. The lack of data in some vital areas;
3. The increasing diversity of data (from images to languages);
4. The vulnerability of data (as the data age or the software to read them become obsolete);
5. The lack of capability in data management;
6. Missed opportunities to collaborate (which is why e-research is so vital, particularly for Australia to enable better use of data and to overcome our relative isolation);
7. Impediments to discover, preserve, share and re-use data (collected and kept in silos so many people collect even more data without knowing what is already available – the need for good longitudinal data for such things as climate change make the discovery of such information vital);
8. Lack of relevant skills (such as mathematical biologists or biological/genetic mathematicians);
9. Lack of global engagement.

The vision which the Working Group had for Australia is:

Australia is managing increasing volumes and complexity of data to enhance our country's scientific, economic and social prosperity and to protect it from threats.

The key data issues which we included in the report were digitisation, capture, preservation, storage, discoverability, integration, interoperability, sharing, re-use, accessibility (for users), security and privacy.

The international community has clear policies about open access and data sharing with OECD (www.oecd.org), ICSU (www.icsu.org) as well as Europe and the Americas having clear philosophies and guidelines encouraging open access and wider use of research data because productivity and quality will be increased. The Australian Bureau of Meteorology summarised it well:

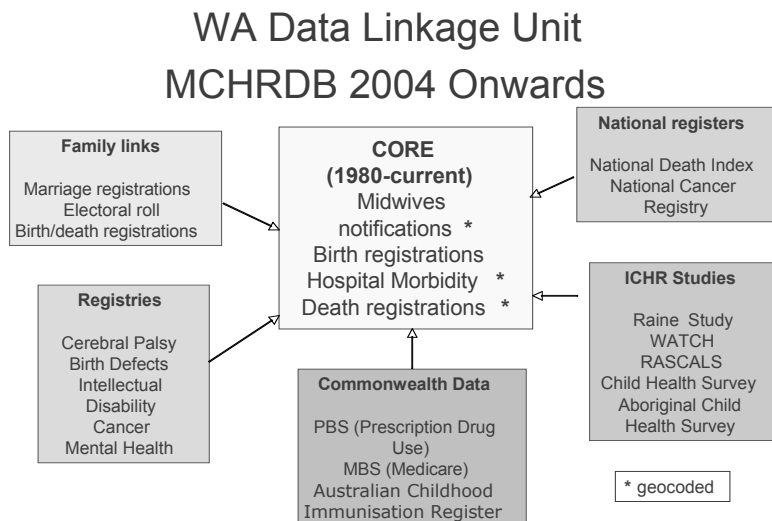
... foregoing proprietary rights to data and making them freely available actually benefits the individual as well as the community at large ...

The recommendations from the report are included as an appendix (Appendix 1) to this chapter. In addition to those about repositories and changing the culture about sharing, access and collaboration, recommendation eight clearly outlined the need to allow researchers to access and link individual data on populations, while protecting privacy.

POPULATION DATA AND RECORD LINKAGE

Record linkage brings together records from different sources relating to the same individual. It is used for administrative purposes, case management and investigation of crimes etc, and also for describing population trends and characteristics in important areas such as the health and well being of the population. It is in the context of maternal, child, adolescent, family and Aboriginal health and development that my group has the most experience in using individual records and linking them together to answer important questions aimed at improving outcomes.

Figure 1



Source: Telethon Institute for Child Health Research

Our interest in obtaining record-linked data began in the late 1970s when there were major concerns about thalidomide and hence other drugs causing birth defects and also whether the introduction of aggressive neonatal intensive care and resuscitation of premature babies would lead to increases in brain damage. We established registers of both birth defects (1980) and cerebral palsies (1977) and then linked them back in to birth registrations and perinatal data to create the Maternal and Child Health Research Data Base.³ Not only did the linkage enable us to study the patterns and causes of all major perinatal and paediatric problems for the whole population, we also confirmed that intensive care did increase cerebral palsy rates in preterm babies, that birth asphyxia was a rare cause of cerebral palsy and commenced

³ F J Stanley et al, 'A population database for maternal and child health research in Western Australia using record linkage' (1994) 8 *Paediatr Perinat Epidemiol* 433–47.

the birth defects studies that led to the confirmation that periconceptual folate prevented spina bifida and related defects.⁴

Figure 1 shows how the original linkage has now grown to include a number of additional databases, with the capacity to link in drugs, immunisations, hospitalisations, cancers and mental health problems. These data on the total population which come from a variety of agencies, statutory and vital statistical collections and special registers and studies provide WA with the one of the most comprehensive tools for monitoring, evaluating and investigating health and health services in the world.⁵ Most of the data are collected for administrative purposes and are brought together for re-use to answer specific questions about disease occurrence or health service quality or effects. A list of projects which have been done using data linkage in WA is available at www.populationhealth.uwa.edu.au/__data/page/63033/projects_1995-2003.pdf.

The advantages of record linkage are:

1. Large sample sizes, no exclusions and hence unbiased data for analysis;
2. Cheap compared with studies which trace individuals, seek consent and collect information directly;
3. Valid and reliable data are obtained on sensitive issues so that individuals are not upset by the research;
4. Survey burden on populations is reduced (particularly relevant to some population subgroups such as Aboriginal who 'have been researched to death' as one of our Aboriginal researchers noted);
5. Fast effective linkage technology is now available;
6. Privacy can be protected by technology and protocols;
7. Better data for policy, planning and evaluation; and
8. Administrative data sets are used and improved.

⁴ C Bower and F J Stanley, 'Periconceptual vitamin supplementation and neural tube defects: evidence from a case-control study in Western Australia and a review of recent publications' (1992) 46 *Journal of Epidemiology and Community Health* 157–61.

⁵ E L Brook et al, Western Australia Data Linkage Unit, Department of Health, *Summary report: research outputs project, WA data linkage unit (1995–2003)*.

Of course we are limited by what is available in the data and privacy issues still need to be addressed.

HOW CONSENT CAN LEAD TO BIAS

Consent is essential for all research involving the direct participation of individuals such as questionnaires, interviews, blood or other tissue sampling and clinical trials. But not all research requires consent and with population data the seeking of consent is neither feasible, cost-effective nor scientifically valid.

Non-participation in surveys where consent is sought comes mostly from an inability to trace the person and rarely from refusal to participate. Most people participating in studies conducted in our institute do so because they are altruistic and wish to help others.⁶ Non-participation is variable, unpredictable and can result in significant bias. Bias is defined as the distortion of the true relationship between exposure and outcome due to flaws in either study design or analysis. People who do not participate in studies can be very different from those who do and hence the analysis can be biased. An example is HIV status where anonymous testing of all blood samples in a population shows much higher rates than when consent is sought. This could result in very different services being developed.

The magnitude and direction of bias from seeking consent in epidemiological studies is unpredictable, not quantifiable, may well explain the differences in risks between studies (e.g. coffee is good for you one day and bad the next!) and provides poor information for health services and epidemiological research.

Examples of significant bias in studies which sought consent include a large meta-analysis of breast cancer and termination of pregnancy and the Canadian stroke register.⁷ In both of these the data obtained from seeking consent and interviewing patients produced such biased data as to be clinically and scientifically useless. The stroke register spent over

⁶ H Bailey et al, 'Applying persuasion principles did not increase questionnaire response: A randomised control trial of a fridge magnet gift' (2007) 14(2) *Australasian Epidemiologist* 6–10.

⁷ V Beral et al, 'Breast cancer and abortion: collaborative reanalysis of data from 53 epidemiological studies, including 83 000 women with breast cancer from 16 countries' (2004) 363(9414) *Lancet* 1007–16; J V Tu et al, 'Impracticability of informed consent in the Registry of the Canadian Stroke Network' [see comment] (2004) 350(14) *N Engl J Med* 1414–21.

\$500 000 over two years on research nurses whose only job was to seek consent from the patients and their carers. A no-consent register would have obtained 100% of patients and been useful to study the natural history of the disease, whether early treatments worked and what their long term impact would have been – all vital questions to help improve the health care system.

HARMONISING PRIVACY AND ACCESS: CAN WE HAVE A WIN:WIN?

In Australia and many other countries, legislation allows access to individual data and record linkage without consent under certain conditions which are governed by ethics committees and medical research funding agency guidelines (see www.nhmrc.gov.au). Ethics committees are guided to allow these activities when it is impractical to seek consent.

Both in Australia and the UK there is a trend towards serving the privacy lobby agenda at the expense of allowing data access for the public good.⁸ This has reduced the amount of record linkage and use of health records for research considerably in UK. One of the main barriers is the interpretations of the legislation by ethics committees. There seems to be a poor public (and ethics committee) awareness of the methods of such research and its value to society.⁹

The WA Data linkage protocols we have developed over many years provide linked databases to researchers who have approval from ethics committees, who have an approved protocol with the Data Linkage Unit committee and approval by the Confidentiality of Health Information Committee. The preparation of the linked data base is done in two stages – the identified information is linked without any of the clinical or sensitive information attached and then the linked information without any identifiers is given to the researchers. Analyses are done on data sets that cannot identify individuals. We believe this is a win:win.

⁸ CW Kelman, A J Bass and C D Holman, 'Research use of linked health data--a best practice protocol' (2002) 26(3) *Australian and New Zealand Journal of Public Health* 251–5.

⁹ Academy of Medical Sciences, *Personal data for public good: using health information in medical research* (2006) <www.acmedsci.ac.uk/images/publication/Personal.pdf>.

FINAL COMMENTS

There have been some recent activities and debates about these issues in UK, USA, Canada and Australia as researchers, policy makers and others realise that the pendulum may have swung too far in favour of privacy. The Australian Law Reform Commission is currently reviewing the privacy legislation and may well help to clarify these issues. The NHMRC is just about to publish a new set of guidelines which outline the Australian code for the responsible conduct of research.

In the UK Academy of Medicine report, Professor John Harris (University of Manchester) asked ‘are patients morally obliged to participate in research projects as a “mandatory contribution to public good,” particularly for those aimed at preventing serious harms and providing important benefits?’

And, at an Australian conference on harmonising privacy and access Professor Eric Meslin commented: ‘[w]e are optimistic that a win:win is possible where privacy is protected, where important health research can proceed. There is evidence that a shift in thinking is possible on behalf of ethics review committees, regulators, researchers and the public.’¹⁰

We believe that we can work towards a win:win. For this to happen we need ethics committees to understand and accept current guidelines which allow identifiable data to be used for research without consent as long as privacy issues are addressed and the rationale is acceptable. We need to develop Australian privacy and research best practice for the researchers using personal data. We need to make the public more aware of how personal records are used and how research is done and why this is of great benefit to society.

I would like to end with a quote from Professor Lawrence O Gostin:

In the late 20th century, scholars and politicians posed a key question. “What desires and needs do you have as an autonomous rights bearing person to privacy, liberty and free enterprise?” Now it is important to ask another kind of question “What kind of community do you want and deserve

¹⁰ Professor Eric Meslin, Indiana University, Centre for Bioethics (2005) personal communication.

to live in, and what personal interests are you willing to forgo to achieve a good and healthy society?”¹¹

APPENDIX ONE – RECOMMENDATIONS FROM PRIME MINISTER'S SCIENCE, ENGINEERING AND INNOVATION COUNCIL (PMSEIC)

A National Strategic Framework for Scientific Data

Recommendation 1

That Australia's government, science, research and business communities establish a nationally supported long-term strategic framework for scientific data management including guiding principles, policies, best practices and infrastructure.

Recommendation 2

That a high level expert committee be established to provide the leadership role in progressing the formation of the long-term strategic framework for scientific data management.

The National Network of Digital Repositories

Recommendation 3

That the necessary policy and programmes be implemented with a view to establishing a sustainable publicly funded network of federated digital repositories.

Recommendation 4

That the expert committee consider the development of a strategic roadmap for the implementation and evolution of the national network of federated digital repositories.

¹¹ L. O. Gostin, 'Law and ethics in population health' (2004) 28(1) *Australian and New Zealand Journal of Public Health* 7–12.

Data Management, Access, Sharing and Collaboration – Changing the Culture

Recommendation 5

That standards-based technologies be adopted and that their use be widely promoted to ensure interoperability between data, metadata, and data management systems, providing authentic users of the data with appropriate processes and safeguards.

Recommendation 6

That the principle of open equitable access to publicly-funded scientific data be adopted wherever possible and that this principle be taken into consideration in the development of data for science and programmes.

As part of this strategy, and to enable current and future data and information resources to be shared, mechanisms to enable the discovery of, and access to, data and information resources must be encouraged.

Recommendation 7

That funding agencies offer incentives to encourage researchers and institutions to:

- Develop data management plans for each research grant application involving data collection and generation, and that standards be made freely available and widely disseminated so as to encourage best practice in data management;
- Introduce policies and practices to encourage collaboration and sharing of data across Australia's scientific research institutions and across agencies;
- Analyse and re-use existing data.

Ensuring there are no Regulatory Impediments

Recommendation 8

That funding agencies such as the NHMRC and ARC ensure that best practices and policies are developed and followed that allow bona fide researchers to access individual population data, including and linking of data from multiple sources, whilst protecting privacy, and ensuring that ethics committees fully understand these policies and their rationale.

Recommendation 9

That in the context of developing the strategic framework for scientific data management, Australia's intellectual property approaches be checked to ensure they do not impede the sharing of data.

In particular, it should take into account the OECD Committee for Scientific and Technological Policy guidelines on access to research data and the International Council for Science statements about the benefits of sharing data.

Skills for Data Management

Recommendation 10

That data management expertise becomes a core skill for researchers, including graduate and postgraduate science students across all disciplines, and that they receive data management training as part of their education.

Recommendation 11

That the Australian Government give early consideration to the finding of the e-Research Coordinating Committee regarding changing research behaviour, practices and skills¹².

¹² Recommendations from R Batterham et al, Prime Minister's Science, Engineering and Innovation Council, Working Group on Data for Science, *From data to wisdom: Pathways to successful data management for Australian Science* (2006) <www.dest.gov.au>. Copyright Commonwealth of Australia reproduced by permission.

PRIVACY REGULATION AND e-RESEARCH

Andrew Hayne¹

INTRODUCTION

The Office of the Privacy Commissioner appreciates the kind invitation from the Law Faculty at Queensland University of Technology to present at the 2007 Legal Framework for e-Research Conference. This legal framework project coincides with a key period for privacy regulation in Australia, most significantly due to the current inquiry into privacy law being conducted by the Australian Law Reform Commission (ALRC). At the same time, public policy is increasingly examining how best to facilitate research interests through the use of personal information. The Office notes, for example, the National Data Network initiative,² as well as the inquiry conducted by the Productivity Commission³ into the role of research in Australia, to which the Office made a submission.⁴

In this chapter I aim to provide a brief overview of federal information privacy regulation, particularly as it applies to health and medical research, as well as to thumbnail possible opportunities for reform that may emerge from the current ALRC inquiry. These opportunities are

¹ Deputy Director of Policy, Office of the Privacy Commissioner.

² <<http://www.nationaldatanetwork.org/ndn/ndnhome.nsf/Home/Home>>.

³ See Public Support for Science and Innovation <<http://www.pc.gov.au/study/science/finalreport/index.html>>.

⁴ The Office of the Privacy Commissioner, *Submission to the Productivity Commission: Research Study into Public Support for Science and Innovation* (2006) <http://www.privacy.gov.au/publications/sub_prod_science072006.html>.

discussed in detail in the Office's submission to that inquiry, available from our website.⁵

OVERVIEW OF THE *PRIVACY ACT* 1988

An important starting point in understanding privacy regulation is to recognise that the *Privacy Act 1988* provides principle-based, technology neutral regulation.

The intention of principle-based law is to emphasise the objectives of the law rather than prescribe what the regulated party may do.

Principle-based law is aimed at encouraging organisations to understand the policy underpinning behind the law and adapt their practices accordingly; not just to prevent intervention from the regulator, but because they recognise the purpose and intent of the law.⁶

Principle-based law also sits comfortably with government policy favouring co-regulation, whereby business is left to pursue solutions that are appropriate to their industry, structure and circumstances, while still meeting the policy objectives of the regulation.

Technological neutrality is intended to recognise the inherent difficulty of keeping statute law up to date with new and emerging technologies.

The Office believes that the *Privacy Act* should continue to be technologically neutral. It is often difficult to envisage how technology will evolve or what new technologies may emerge. It would therefore be extremely difficult to respond effectively to dynamic technological development.⁷

At the same time, to accommodate particular emerging technologies that may create privacy risks, the Office has proposed to the ALRC inquiry that the *Privacy Act* should provide the flexibility for the Privacy

⁵ The Office of the Privacy Commissioner, *Submission to the Australian Law Reform Commission's Review of Privacy – Issues Paper 31* (2007) <<http://www.privacy.gov.au/publications/alrc280207.html>>.

⁶ See also Karen Curtis (Privacy Commissioner) 'Reducing overlap, duplication and inconsistency' (Speech delivered at the Australian Regulatory Reform Evolution 2006, Canberra, 24 October 2006) <http://www.privacy.gov.au/news/speeches/sp05_06.pdf>.

⁷ This is discussed in further detail in the Office of the Privacy Commissioner, *Submission to the Australian Law Reform Commission's Review of Privacy – Issues Paper 31* (2007) Chapter 11 <<http://www.privacy.gov.au/publications/submissions/alrc/c11.html#L25052>>.

Commissioner, subject to Parliamentary oversight, to make binding codes that go to specific acts or practices that may be enabled by new or emerging technologies.⁸

MEANING OF ‘PERSONAL INFORMATION’

It is important to recognise that the *Privacy Act* focuses its regulatory functions on information privacy. In turn, the scope of information privacy is determined by the meaning of ‘personal information’.

The statutory definition of personal information is contextual, in that it refers to information or opinion about an individual whose identity is apparent *or can be reasonably ascertained*. Clearly, whether an identity can be reasonably ascertained will be determined by the context in which that information is held, including the availability of technologies that may reasonably *re-identify* information that is putatively *de-identified*.

For example, Robert Gelman, in *Public Record Usage in the United States*,⁹ cites research that reveals:

... the Cambridge, Massachusetts voter registration list has 55,000 voters. Twelve percent of voters have unique birthdates. So if a person of voting age lives in Cambridge, the voter might be identified just from the birthdate on the voter list. With birthdate and gender, 20% of voters are unique. With birthdate and five-digit zip code, 69% are unique. With birthdate and nine-digit zip code, 97% are unique. More broadly, 87% of Americans can be identified just by birthdate, five digit zip code, and gender.

More recently, the Office notes the widely publicised case whereby 20 million putatively de-identified internet search records on 650 000 AOL users were made publicly available. By examining linkages between different searches, a *New York Times* journalist found that:

⁸ This is discussed in the Office of the Privacy Commissioner, *Submission to the Australian Law Reform Commission’s Review of Privacy – Issues Paper 31* (2007) Chapter 6 and Chapter 11 <<http://www.privacy.gov.au/publications/alrc280207.html>>.

⁹ Available at <http://www.cnil.fr/conference2001/eng/contribution/gellman_contrib.html>.

It did not take much investigating to follow that data trail to Thelma Arnold, a 62-year-old widow who lives in Lilburn, Ga.¹⁰

In the view of the Office, this contextual element is one of the strengths of the definition, allowing it to respond to change and technological advance, as well as the particulars of a given context. In order to alleviate any confusion generated by the flexibility of the term, the Office intends to issue further guidance material on the meaning of 'personal information' in a regulatory context.¹¹

STATUS OF 'HEALTH INFORMATION' IN PRIVACY REGULATION

The *Privacy Act* also deals expressly with health information, which is defined in broad terms and exists as a subset of personal information.¹²

Consistent with the second reading speech for the Privacy Amendment (Private Sector) Bill 2000, the community expects that such health information will be afforded privacy protections that are in addition to those applying to non-health information.

In the second reading speech for that Bill, the then Attorney-General, the Hon Daryl Williams QC, said that:

The government recognises that the Australian public considers their health records to be particularly sensitive ... The bill provides additional protections in relation to the use and disclosure of health information, as such information is

¹⁰ M Barbaro and T Zeller 'A Face Is Exposed for AOL Searcher No. 4417749', *New York Times*, 9 August 2006, <http://www.nytimes.com/2006/08/09/technology/09aol.html?ei=5087&en=fc3fb3310bf58bd7&cex=1171771200&excamp=mkt_at1&pagewanted=all>.

¹¹ The adequacy of the definition of 'personal information' is discussed in the Office of the Privacy Commissioner, *Submission to the Australian Law Reform Commission's Review of Privacy – Issues Paper 31* (2007) Chapter 3 <<http://www.privacy.gov.au/publications/submissions/alrc/c3.html#Personal>>.

¹² 'Health information' is discussed further in the Office of the Federal Privacy Commissioner, *Guidelines on Privacy in the Private Health Sector* (2001) A.3.2 <http://www.privacy.gov.au/publications/hg_01.html#a32>.

clearly considered by the community to be particularly sensitive.¹³

REGULATION AFFORDED BY THE *PRIVACY ACT*

In regard to the *Privacy Act's* jurisdiction, the *Act* sets out 11 principles, called the Information Privacy Principles, which apply to most Australian Government agencies, and 10 principles, termed the National Privacy Principles, which apply to all private sector bodies with turnover greater than \$3 million, as well as to all health service providers in the private sector.

Significantly, neither set of privacy principles apply to state agencies, including public health systems, nor to most public universities, except where established under Commonwealth law.

The two sets of principles, while having differences in a number of areas, share underlying objectives, including ensuring that individuals know who has personal information about them, what will be done with it and that it will be handled with appropriate security.

Common to both principles is the general requirement that personal information, including health information, should only be used or disclosed for the purpose for which it was initially collected, unless an exception specified in the *Privacy Act* applies - I will return briefly to these exceptions and secondary purposes shortly.

However, notwithstanding the commonalities between the principles, the Office is of the view that maintaining two sets of privacy principles causes unnecessary complexity for all stakeholders. Law reform could, and should, usefully include amendment to create a single set of privacy principles.¹⁴

¹³ The Hon Daryl Williams QC, Second Reading Speech Privacy Amendment (Private Sector) Bill 2000 <http://parlinfoweb.aph.gov.au/piweb/TranslateWIPILink.aspx?Folder=HANSARDR&Criteria=DOC_DATE:2000-11-08%3BSEQ_NUM:8%3B>.

¹⁴ The proposal for a single set of privacy principles is discussed throughout the Office of the Privacy Commissioner, *Submission to the Australian Law Reform Commission's Review of Privacy – Issues Paper 31* (2007), though most directly in Chapter 4 <<http://www.privacy.gov.au/publications/submissions/alrc/c4.html>>.

OFFICE VIEW ON MEDICAL RESEARCH

The Office recognises that there is an important social interest in enabling medical researchers to have access to health information in certain circumstances. The *Privacy Act* is not intended to restrict important medical research. While health information, being sensitive information, is afforded extra protection under NPPs, the *Privacy Act* recognises the desirability of health and medical research by enabling health information to be collected, used and disclosed for these purposes, in some cases, without consent.

FUNCTIONS OF THE PRIVACY COMMISSIONER

It is useful to note that the Privacy Commissioner has express functions under the *Privacy Act* concerning health and medical research. Most significantly, these functions include to approve Guidelines made by the National Health and Medical Research Council (NHMRC) under sections 95 and 95A of the *Privacy Act*.¹⁵ These guidelines provide a framework for non-consensual research, and I will return to them shortly.

The Privacy Commissioner also receives reports from the NHMRC on the operation of these Guidelines. These reports serve important oversight functions for the operation of the guidelines. To promote transparency, the Office can see merit in progressing to a point where these reports are made publicly available.

QUESTIONS FOR THE RESEARCHER

For the researcher, the application of the *Privacy Act* first turns on whether the data involved meets the definition of personal information. If it does not, then neither set of privacy principles apply.

This again raises the issue of what does personal information mean – whether or not data satisfies the statutory definition will depend on the circumstances in which it is held and, crucially, whether an individual's identity is apparent or reasonably ascertainable.

¹⁵ The section 95 and 95A Guidelines are available at
<<http://www.privacy.gov.au/health/guidelines/index.html#2>>.

If it is established that research data is regulated as personal information, the *Privacy Act* offers a number of mechanisms by which it may be handled for research purposes. These include where the information was initially collected for the primary purpose of conducting that research project.

The Office of the Privacy Commissioner also recognises that the use of personal information for the *secondary* purpose of research is of significance to researchers. It is perhaps most common for health information, in particular, to be collected for purposes other than research, such as the clinical care of the individual. Nonetheless, this information may be of considerable value in a research context.

Further, as shared electronic health records systems evolve, there would seem every chance that richer repositories of health information may emerge.¹⁶

The *Privacy Act* provides various mechanisms by which health information may be used for the secondary purpose of research.

For example, this may occur with the consent of the individual; in this regard, researchers may usefully bear in mind that consent may be express or implied, and may be written or verbal.

The Parliament, in recognition of the important role of health and medical research, has also acknowledged that, in some circumstances, health information should be available for important research activities where it is impracticable to gain the individual's consent.

The Office has issued guidance material explaining that impracticability may include where:¹⁷

¹⁶ The issue of electronic health records is discussed at question 8–5 of the Office of the Privacy Commissioner, *Submission to the Australian Law Reform Commission's Review of Privacy – Issues Paper 31* (2007) <<http://www.privacy.gov.au/publications/submissions/alrc/c8.html#L20635>>. The Office has also discussed its views on EHRs more generally in submissions to the former HealthConnect project office, see Office of the Privacy Commissioner, *Submission on the HealthConnect Business Architecture* (2005 Version 1.9) <<http://www.privacy.gov.au/publications/hlthcnctsub.pdf>>.

¹⁷ The question of when consent may be impracticable is discussed at question 8–30 of the Office of the Privacy Commissioner, *Submission to the Australian Law Reform Commission's Review of Privacy – Issues Paper 31* (2007) <<http://www.privacy.gov.au/publications/submissions/alrc/c8.html#L22503>>.

- individuals may be uncontactable due to there only being old records available;
- the individuals of interest may be part of a demographic group that is typically difficult to contact, including remote, transient or indigenous groups;
- the sheer number of records involved may cause excessive logistical problems; and
- where seeking consent may in itself fundamentally and unavoidably undermine the integrity of the research methodology.

This mechanism for non-consensual research is facilitated through sections 95 and 95A of the *Privacy Act*.¹⁸ These sections apply, respectively, to Commonwealth agencies and to private sector organisations.

These sections require the NHMRC to make guidelines, approved by the Privacy Commissioner, setting out under what circumstances non-consensual research may proceed.

The guidelines provide a framework to ensure privacy protection of health information that is collected or used or disclosed in the conduct of research. Under the guidelines, Human Research Ethics Committees (HRECs) are required to approve research, including by considering the affect on the privacy of the research subject.

THE NEED TO HARMONISE SECTIONS 95 & 95 OF THE *PRIVACY ACT*

While the Office broadly supports this form of mechanism, it is apparent that, while having similar policy objectives, sections 95 and 95A display a number of inconsistencies. Agencies, for example, may handle any form of personal information for the purpose of medical research, while organisations are limited to handling ‘health information’ albeit for apparently much broader purposes of ‘research relevant to public health or public safety’. This would appear to limit, for example,

¹⁸ Sections 95 and 95A are available at <[http://www.comlaw.gov.au/ComLaw/Legislation/ActCompilation1.nsf/previewlodgmentattachments/409069FCABD20271CA25725C008385B5/\\$file/Privacy1988_WD02HYP.htm#param220](http://www.comlaw.gov.au/ComLaw/Legislation/ActCompilation1.nsf/previewlodgmentattachments/409069FCABD20271CA25725C008385B5/$file/Privacy1988_WD02HYP.htm#param220)>.

the linking of health information with non-health information, notwithstanding that such linkages may be for public health or safety research.

Stakeholders have previously expressed the view that the existence of two sets of Guidelines regulating the public and private sectors was causing difficulties for researchers and ethics approval processes.¹⁹

The differing requirements of Sections 95 and 95A are inconsistent and confusing. Accordingly, in our recent submission to the ALRC inquiry, the Office has pointed to the potential benefits of a simplified framework for the regulation of how personal information may be handled, without consent, for health related research by organisations and agencies.

REVIEWING THE *PRIVACY ACT*

As I have mentioned already, the reform of privacy law is very much a live matter, and may have significant implications for research.

Since 2003, there have been three reviews instigated of the current state of federal privacy regulation in Australia, albeit with different objectives and terms of reference.

The reviews conducted by the Office of the Privacy Commissioner and the Senate Legal and Constitutional Affairs Committee have led up to the current inquiry by the ALRC.

WHAT MIGHT REFORM OFFER MEDICAL RESEARCH?

An important question is to ask, what might useful law reform look like, particularly as it affects health and medical research?

Retain Strong Protections for Health Information

The Office of the Privacy Commissioner would expect that any such reform should proceed from the recognition of the importance individuals place on how their health information is handled.

¹⁹ This, and other research related issues, was discussed in the Office of the Privacy Commissioner, *Getting in on the Act: The Review of the Private Sector Provisions of the Privacy Act 1988* (2005) <http://www.privacy.gov.au/act/review/review2005.htm#7_3>.

Individuals' engagement with the health sector remains largely premised on the assumption that they can rely on providers to maintain the privacy and confidentiality of their health information. Drawing on the World Medical Association's recent 2006 Declaration of Geneva, providers assert that they will 'respect the secrets that are confided in me, even after the patient has died'.

The Office believes that codified privacy regulation, which seeks to balance the public interest in privacy, with the public interest in health and medical research, plays an important role in sustaining community confidence about how health information may be used for research purposes.²⁰

From this basis, the Office supports the ongoing role of HRECs as providing appropriate institutional oversight of human research.²¹

While submissions to the Office's 2003 review referred to concerns about the adequacy of HREC resources, and whether HREC decision making may, on occasion, be unnecessarily conservative in regard to privacy, the Office remains of the view that the existence of institutional ethical oversight has served Australia effectively and promoted community confidence that abuses committed in the name of research in other countries, are unlikely to happen here.

Harmonise the Section 95 and 95A Mechanisms

At the same time, the Office has proposed that simplifying and harmonising the section 95 and 95A processes, including by making a single, common set of guidelines for Commonwealth agencies and the private sector, would assist HREC decision making by reducing unnecessary complexity.²² The Office has already committed to work

²⁰ This theme is also discussed in the Office of the Privacy Commissioner, *Research Study into Public Support for Science and Innovation: Submission to the Productivity Commission* (2006) <http://www.privacy.gov.au/publications/sub_prod_science072006.html>.

²¹ The role of HRECs in providing institutional oversight of research is discussed at question 8–31 of the Office of the Privacy Commissioner, *Submission to the Australian Law Reform Commission's Review of Privacy – Issues Paper 31* (2007) <<http://www.privacy.gov.au/publications/submissions/alrc/c8.html#L22607>>.

²² The question of harmonising the section 95 and 95A mechanisms is discussed in detail at question 8–32 of the Office of the Privacy Commissioner, *Submission to the Australian Law Reform Commission's Review of Privacy – Issues Paper 31* (2007) <<http://www.privacy.gov.au/publications/submissions/alrc/c8.html#L22695>>.

with the NHMRC to explore ways to simplify reporting obligations faced by HRECs.

One Set of Privacy Principles

More generally, the proposal for a single set of privacy principles, common to agencies and organisations, would similarly lessen regulatory confusion as to how research may be undertaken.

Clarify Interaction with State and Territory Law

The Office has also proposed reform to the *Privacy Act* to remove any uncertainty as to the role of State and Territory privacy laws to the private health sector.²³

In this regard, the Office has previously stated that the best advice available to it is that where an act or practice is regulated by the Commonwealth *Privacy Act*, then it is not regulated by a State or Territory privacy Act. On this basis, the State and Territory health privacy Acts are restricted in their application to the relevant State or Territory public sector.

Equally though, the Office has recognised that the matter is not fully settled and that other parties may have differing advice. The Office's view is that this lack of certainty creates a major potential obstacle to effective and consistent privacy regulation in the Australian federal system.

The Office has proposed that amending the *Privacy Act* to make clear that its provisions 'cover the field' for the regulation of private sector health service providers would be a significant step toward reducing possible uncertainty for those bodies, including in research contexts.

Issues for e-Research

An issue that may have particular import for the e-research agenda include ensuring clarity and certainty around the meaning of 'personal information', particularly in light of the contextual element introduced in its definition through reference to someone's identity being 'reasonably

²³ See question 8–2 of the Office of the Privacy Commissioner, *Submission to the Australian Law Reform Commission's Review of Privacy – Issues Paper 31* (2007) <<http://www.privacy.gov.au/publications/submissions/alrc/c8.html#L20540>>.

ascertainable'. The Office has committed to providing further guidance on this issue.²⁴

Perhaps also significant are the provisions regulating transborder dataflows of personal information. Advances in information technology have allowed information to be sent across the world with speed and efficiency. With the advent of inexpensive high-speed internet connections and the growth of the global economy, Australian agencies and organisations are increasingly operating across national borders. This will equally apply to researchers.

Currently, personal information may only be sent overseas subject to the requirements of National Privacy Principle 9, which include that such transfers should occur where comparable privacy protections apply, either in law or by other agreement, or where the individual consents.

Further analysis may be required to flesh out the privacy law obligations involved in exchanging personal information across borders for research, particularly in regard to such matters as ensuring legal compliance and the role of HRECs in an international context.

The question of how best to regulate datasets established for broad research purposes, such as health registers, remains an important one.²⁵

The Office has noted that many such registers have benefited from the certainty of being established under state or territory law, or on the basis of individual consent.

The Office notes that, with the expansion of electronic health records, it may become increasingly difficult to quarantine research registers from other health information systems. The move towards electronic health records may put increasing pressure on health records to be multi-functional, where they are used for patient-care, as well as epidemiological and other research objectives.

The role of consent in the context of multi-purpose data registers seems unclear, particularly where it may not be known what specific research

²⁴ The meaning of 'personal information' was discussed in the Office of the Privacy Commissioner, *Getting in on the Act: The Review of the Private Sector Provisions of the Privacy Act 1988* (2005) <<http://www.privacy.gov.au/act/review/review2005.htm#8>>.

²⁵ Health registers and datalinkage are discussed at question 8–33 of the Office of the Privacy Commissioner, *Submission to the Australian Law Reform Commission's Review of Privacy – Issues Paper 31* (2007) <<http://www.privacy.gov.au/publications/submissions/alrc/c8.html#1.22811>>.

will be undertaken in the future and, therefore, individuals may not be adequately informed so as to offer truly meaningful and valid consent.

The Office sees merit, therefore, in specific legislative provision being made for the establishment of health data registers that are intended to serve broad research objectives. Doing so would recognise both the value of such registers, and the sensitivity of the information they contain and would offer the certainty, parliamentary oversight and scrutiny needed to sustain community confidence.

CONCLUSION: GOOD PRIVACY SUPPORTS GOOD RESEARCH

In closing, the Office is well aware of criticisms from some stakeholders that privacy regulation unreasonably impedes research in some contexts. The Office believes that regulatory reform to promote simplicity and overcome regulatory uncertainty would likely address many of these concerns.

More generally, though, far from being an obstructing factor, in the Office's view, privacy regulation is a necessary and supporting condition for serving the public interest in the benefits of research. The relationship of trust between health service providers and individuals is vital for sustaining public confidence in the health sector, their participation in effective treatment and the resulting quality of medical research.

A PRIMER IN THE POLITICS OF PRIVACY AND RESEARCH

David Ruschena¹

INTRODUCTION

Privacy legislation in Australia is experiencing interesting times. The Australian Law Reform Commission (ALRC) has published its Discussion Paper on its Review of Australian Privacy Law dealing with the potential amendment of the Commonwealth *Privacy Act*.² The ALRC review is the third review of the *Act* in the past three years, with additional reviews being performed by the Australian Office of the Privacy Commissioner (OPC).³ In short, this is an area where considerable developments are being proposed and considered.

The Discussion Paper recommends significant changes both to the structure of Australian privacy legislation⁴ and to the substantive obligations. Recommended changes to substantive obligations will impact research by imposing obligations regarding the privacy of deceased persons⁵ and third parties whose information was not

¹ Consultant, Health Legal, PhD Candidate, University of Melbourne Law School. This chapter is based on a presentation given at the Legal Framework for e-Research Conference held 11 and 12 July 2007, Gold Coast, by the Queensland University of Technology Faculty of Law. The views expressed in this article are those of the author alone and do not necessarily reflect those of Health Legal.

² Australian Law Reform Commission, Review of Australian Privacy Law, Discussion Paper (September 2007 <http://www.austlii.edu.au/au/other/alrc/publications/dp/72/> at 29 June 2008 (hereafter Discussion Paper).

³ Office of the Privacy Commissioner, *Getting in on the Act: The Review of the Private Sector Provisions of the Privacy Act 1988* (2005).

⁴ The major changes proposed involve creating a single set of principles that would apply to both public and private sector organisations (see Discussion Paper, op cit n 2, Part D), and that federal privacy legislation override state and territory privacy legislation in relation to health services provided by the private sector (ibid, Part H).

⁵ Ibid, Chapter 3.

solicited,⁶ and by imposing an obligation to ensure that the personal information they collect is relevant to the purpose for which it is collected.⁷

These developments should be taken seriously. Some researchers consider that the *Privacy Act* presents a significant obstacle in the conduct of research.⁸ There is no doubt that, compared to untrammelled rights of access, the *Privacy Act* has resulted in higher research costs, lost opportunities, less effective research and sub-optimal quality of data. However, privacy legislation has the potential to be even more burdensome than it is, or even to prevent research from occurring. It is in researchers' best interests to understand how that might occur. These developments are important not just because they might have a chilling effect on research, but because they show that community acceptance of research – and researcher's need to use personal information to obtain significant results – cannot be taken for granted.

The purpose of this chapter is to consider the political and legal landscape that surrounds privacy legislation and to argue that without a commitment by researchers to engage with the Australian society, privacy legislation will remain subject to change in this way.

The chapter will commence by conducting a brief tour of the politics of rights. Privacy legislation was enacted to meet a perceived need, and that perception is more important than the reality. The chapter will then examine how research takes place in accordance with privacy legislation. It is argued that, although research may occur without obtaining the consent of subjects, the exceptions are both less available than they are perceived to be, and do not advance the cause of research generally. Ultimately, however, the framework of privacy law itself provides researchers with significant opportunities to influence the regulatory environment within which they must operate. This can be done in a

⁶ Ibid, Chapter 18.

⁷ Ibid, Chapter 24.

⁸ See, for example, Professor Fiona Stanley 'Record Linkage – Public Good or Invasion of Privacy?' (Paper presented at the 25th International Conference of Data Protection and Privacy Commissioners, Sydney Australia, 10 September 2003). See also the submission by Dr Richie Gun, Department of Public Health, University of Adelaide, *Submission to the Office of the Privacy Commissioner Review of the Private Sector Provisions of the Privacy Act 1988*, 21 December 2004.

simple way: by adopting the rule-of-thumb that wherever consent can be obtained, it should be obtained.

UNDERSTANDING THE POLITICS OF RIGHTS

The *Privacy Act*, and privacy legislation generally, protect information (whether true or not) from which a person's identity can reasonably be ascertained. This type of information is known as 'personal information'.

The Commonwealth *Privacy Act* was enacted following the demise of the 'Australia Card' proposal in 1988. That proposal failed due to significant public opposition to the idea of government collecting and controlling data relating to the activities of Australian citizens. The *Act* was created to meet a perceived need that privacy was a valid right, one that was endangered (whether actually or potentially) and one that needed to be protected.

Privacy legislation therefore exists within the politics of rights. By this phrase, I mean the discourse within society about what interests are worth protecting and how strong that protection should be. Inevitably, the politics of rights involves questions of balance. For example, the right to free speech must be balanced against the right not to be vilified because of one's gender, sexual orientation, race or religion. The politics of any given right is a contest between differing views about when a generally acceptable balance is reached. Most importantly, the acceptability of any given balance can change. For example, Australian society generally acknowledges the right of individuals to seek compensation for personal injury, if that injury is caused by another person. Nevertheless, that right was severely curtailed during the recent tort 'reforms', because of the perception that compensation was being awarded too easily, in amounts that were excessive and in respect of losses that were properly the plaintiff's personal responsibility.⁹

The best example of how the politics of rights may be used is the actions of the tobacco industry. Fifty years after evidence began to accumulate that showed the link between cigarette smoking and cancer, the tobacco industry is still able to function effectively in the manufacture and sale of

⁹ D Ipp, Department of Treasury, Canberra (2002) *Review of the Law of Negligence*.

their products.¹⁰ Despite killing almost 20 000 Australians per year,¹¹ the tobacco industry's actions go on almost completely unregulated.¹² This has occurred because the tobacco industry has been very effective in controlling the terms of the debate. Instead of concentrating on the public health, the tobacco industry has focussed on 'the individual's right to choose'. Instead of allowing the debate to be about the cost of healthcare, the tobacco industry has focussed on the perils of taxation and government bureaucracy.¹³ Lastly, whenever regulation of the industry is proposed, the industry has suggested that this is simply a precursor to prohibition, which it blandly states 'doesn't work'.

What are the lessons for privacy in the politics of rights? First, simple messages with emotional appeal are very powerful. Prohibition and freedom are simple messages; so is 'Big Brother'. Second, perception is more important than reality. If people believe that something is wrong, they will support measures to change it, even if nothing is 'really' wrong. Third, and most important, just because you are 'right' doesn't mean you will succeed in achieving your goals.

It is therefore important to remember that all it takes is for a 'privacy scare' to occur, and politicians may see that it is in their best interests electorally to alter the privacy legislation to better protect the right of privacy. Politicians tend to infer society's views on an issue from how that issue plays in the media.¹⁴ If the media can frame a message in the right way, and this generates traction within the community, politicians will be tempted to introduce legislation to meet the perceived demand.¹⁵

¹⁰ See R Kluger, *Ashes to Ashes: America's Hundred-Year Cigarette War, the Public Health and the Unabashed Triumph of Philip Morris* (1996); R A Glantz and E D Blaback, *Tobacco War* (2000).

¹¹ B Ridolfo and C Stevenson, Australian Institute of Health and Welfare, (2001) *The quantification of drug-caused mortality and morbidity in Australia, 1998* (Drug Statistics Series No 7) AIHW category no PHE29.

¹² N Gray 'The Modern Cigarette, an Unregulated Disaster' (2007) 187 (9) *Medical Journal of Australia* 502.

¹³ Glantz S A & Balbach E D *Tobacco War: Inside the California Battles* (2000).

¹⁴ Yanovitzky I 'Effects of News Coverage on Policy Attention and Actions: A Closer Look at the Media – Policy Connection' (2002) 29 (4) *Communication Research* 422.

¹⁵ Kingdon R *Agendas, alternatives and public policies* (1984).

Ultimately, researchers should acknowledge that all their actions fashion the regulatory environment in which they operate, not just those that are explicitly political.

WHAT DOES THE PUBLIC ACTUALLY THINK?

Many researchers appear to take the stance that the OPC is part of a 'privacy lobby',¹⁶ which stands in the way of progressive and necessary research; research the Australian society understands and with which it generally agrees. The attitude of the 'privacy lobby' is said not to be representative of the Australian public. The reality is that the stance taken by the OPC simply reflects their understanding of the attitudes of Australian society.¹⁷ The relationship of those attitudes to privacy and research is somewhat ambivalent. Surveys commissioned by the OPC, the Commonwealth Department of Health and Ageing (DoHA) and by the National Health and Medical Research Council (NHRMC) report that anywhere between one third and two thirds of Australians are against the use of identifying information for research without the consent of the subject.¹⁸ One fifth of individuals in one survey reported reluctance to provide their medical history or health information to *any* organisation.¹⁹

These attitudes exist even where research is accepted as important. In one survey, although 83% of respondents believed such research was critically or very important, 73% of respondents believed it was critically or very important to get consent for each research study.²⁰ Given such attitudes, the representations made by the OPC are at least as

¹⁶ Presentation of Professor F Stanley at the Legal Framework for e-Research Conference, 12 July 2007, Gold Coast.

¹⁷ See Office of the Privacy Commissioner, *Submission to the Australian Law Reform Commission's Review of Privacy - Issues Paper 31* (2007) [11], [41], [233], [345] etc.

¹⁸ See Office of the Privacy Commissioner, *Getting in on the Act: The Review of the Private Sector Provisions of the Privacy Act 1988* (2005) 210–11; National Health and Medical Research Council, *Submission to the Office of the Privacy Commissioner Review of the Private Sector Provisions of the Privacy Act 1988*, 10 December 2004.

¹⁹ See Office of the Privacy Commissioner, *Community Attitudes to Privacy* (2004) 6.2 <<http://www.privacy.gov.au/publications/rcommunity>>.

²⁰ L Damschroder, 'Patients, privacy and trust: Patients' willingness to allow researchers to access their medical records' (2007) 64 *Social Science and Medicine* 223–35.

representative of the Australian community as those made by the researchers.

The best way to counter the representations is to address the community's concerns.

CONDUCTING RESEARCH THAT COMPLIES WITH PRIVACY LEGISLATION

The *Privacy Act* allows the collection, use and disclosure of health information that identifies an individual in five relevant circumstances:

1. with the individual's consent;²¹
2. for a secondary purpose that is:
 - 2.1. directly related to the primary purpose; and
 - 2.2. within the individual's reasonable expectations,²²
3. for research relevant to public health or public safety;
4. for the compilation or analysis of statistics relevant to public health or public safety; or
5. for the management, funding or monitoring of a health service.²³

Obtaining the subject's consent can be inconvenient, difficult or impossible, depending on the size, timing, subject matter, importance and methodology of the research proposal. The population which forms the subject of the study may be dead, transient, remote or simply uncooperative. In many cases, researchers may consider that if consent is required, it is simply not worthwhile to perform the study. In such circumstances, many researchers will be tempted to try to bring their research within the exceptions to consent. However, there are a number of traps that mean that care should be taken.

Secondary Purposes

The secondary purpose must be within the reasonable expectations of the individual who forms the subject of the research, not the expectations of the Human Research Ethics Committees (HREC) or the

²¹ *Privacy Act 1988* (Cth) Sch 3, National Privacy Principle 10.1.

²² *Privacy Act 1988* (Cth) Sch 3, National Privacy Principle 10.2.

²³ *Privacy Act 1988* (Cth) Sch 3, National Privacy Principle 10.3.

researcher. Researchers and HREC members form a relatively small subset of the population; one which is aware of the various uses to which health information could be put. As a rule of thumb, the rest of society can only ‘reasonably’ be considered to know what they have seen on television. For example, researchers and HRECs are aware that clinical review of a particular individual may take place within a health service, at conferences, or in the course of multi-site research into the effectiveness of a particular medical technology. By comparison, the rest of society probably only understands clinical review of a particular individual to take place within the health service, and probably only if an adverse event has occurred.²⁴

HRECs, the Public Interest and Impracticability

The last three possible circumstances may only take place where other circumstances exist. These circumstances are where:

- the purpose cannot be served by the collection or use of de-identified information; and
- it is impracticable for the organisation to seek the individual’s consent to the collection; and
- the information is collected:
 - as required by law (other than the *Privacy Act* itself); or
 - in accordance with rules established by competent health or medical bodies that deal with obligations of professional confidentiality which bind the organisation; or
 - in accordance with guidelines approved by the OPC.

The Guidelines published by the OPC provide that where collection, use or disclosure takes place for the purpose of research, that research must

²⁴ Given the uncertainty associated with what the Australian population knows or perceives about research, e-Research and health research, there is an obvious need for further research. Hard research – as always – prevents speculation about the public’s views, which would necessarily be influenced by the speculator’s own perception of the importance of privacy. However, in the absence of such research, it is in researchers’ long term interests to adopt a conservative approach.

be approved by a HREC. HRECs may approve such research only if they consider that the public interest in the research substantially outweighs the public interest in maintaining the level of privacy protection provided by the privacy legislation. This occurs by taking into account:

- the value and public importance of the research;
- the likely benefits to the participants;
- whether the research design can be modified;
- the financial costs of not proceeding with the research;
- the type of personal information being sought;
- the risk of harm to individuals; and
- the extent of a possible breach of privacy.²⁵

While the potential benefits of the research are often apparent, the potential detriment is not always clear. The detriment is, simply, that people will not trust their doctors. Without an assurance that their health information will remain private, people may not seek the health care they need, or may not provide a full medical history. This may in turn increase the risks to their own health and the health of others.²⁶ For example, the Cancer Council has recently stated that some cancer patients are too scared and embarrassed to seek help for their condition because of the stigma of smoking.²⁷ Again, HRECs should remember that they are part of a small portion of society that has both regular contact with doctors and a good understanding of the research that is conducted to supplement medicine.

Lastly, many researchers and doctors consider that the question of impracticability is one to be decided by the HREC. This is incorrect. Impracticability is a question of law. In the opinion of the OPC and the National Health and Medical Research Council (NHRMC), impracticability occurs where:

²⁵ National Health and Medical Research Council, *Guidelines under Section 95 of the Privacy Act 1988* (2000) Guideline 3.2.

²⁶ Australian Government Department of Health and Ageing, *Submission to the Office of the Privacy Commissioner Review of the Private Sector Provisions of the Privacy Act 1988*, 10 December 2004.

²⁷ K Sikora, 'Lung cancer our hidden killer', *Daily Telegraph*, 12 July 2007.

- the subjects are uncontactable due to death or relocation or part of a demographic group that is typically difficult to contact,
- the sheer number of records involved may cause logistical problems,
- the procedures required to obtain consent are likely either to cause unnecessary anxiety for those whose consent would be sought; or
- the objective of the investigation may need to be concealed from subjects in order to minimise various forms of bias.²⁸

Note that this is a high barrier to overcome. Researchers and HRECs should be careful of being seen as too willing to allow research to occur without obtaining consent from the subjects of the research.

A strict reading of privacy legislation, therefore, imposes significant obstacles in the path of researchers seeking to use personal information. This is the intention behind the legislation: to make consent the ‘default’ option which researchers should consider first before seeking to utilise the exceptions set out above.

CONSENT AND ITS BENEFICIARIES

The short term goals of any researcher are the successful completion and publication of research. For health researchers, a long term goal is to promote health research as an essential part of ensuring that people get the best care possible. In such circumstances, obtaining consent directly from subjects (and not relying on the exemptions discussed above) has two significant benefits.

First, obtaining consent is the surest way to avoid litigation challenging research that takes place under an exception to the consent requirement. Such litigation will take place in the courts, and judges are more likely to be sympathetic to the discourse of rights than they will be to postulated future benefits of research. Court decisions are often published in the media, and may result in the imposition of financial penalties. Court decisions also serve to mobilise political forces, by making people think

²⁸ See Information Sheet 9: *Handling Health Information for Research and Management* <http://www.privacy.gov.au/publications/IS9_01.html>; National Health and Medical Research Council, *National Statement on Ethical Conduct in Research Involving Humans* (1999) [14.4].

about social interactions in new ways.²⁹ Adverse judicial decisions therefore pose the most significant risk of a chilling effect on the conduct of research.

Second, the process of obtaining consent from a subject necessarily involves education. To provide informed consent, the subject must be made aware of the research that is occurring and the need for that subject's participation in the research. This normalises research, and participation in research. Every person who is involved in research and who does not experience an intrusion into their privacy will be less taken in by the simple messages, like 'Big Brother', that may be used to undermine research. The longer the process of normalisation goes on, the less public concern will be reflected by the OPC and the rest of the 'privacy lobby'. In other words, one of the primary beneficiaries of the consent process will be the research community generally. The assumption that privacy legislation is something to be tolerated is false, and is a dangerous assumption to make.

CONCLUSION

The ultimate message of this chapter is that regulatory regimes are not set in stone; the ALRC Discussion Paper is perfect evidence of the dynamic of review, discussion and amendment. Even with its frustrations, the current privacy regime cannot be taken for granted. Legislative change is as capable of making research harder as it is of making research easier.

Given this reality, researchers should consider how their actions will affect the long term viability of their discipline. Researchers should adopt the approach that, wherever consent can be obtained, it must be obtained. This approach is demanded by the law, but it also provides researchers with an opportunity to educate the public about the importance of health research. Simply by undertaking the process of education, researchers generate trust, because they show society that society's views are taken seriously.³⁰ In the long term, this is the best way to ensure that research remains viable and legitimate.

²⁹ See McCann M, *Rights at Work* (1992)

³⁰ Tyler T, *Why People Obey the Law* (1992)

Researchers living a one-grant-to-the-next reality may find these comments utopian. I acknowledge that taking a conservative approach to consent might reduce the possibility of significant research being undertaken. However, I am only advocating a general approach, and exemptions – though rare – may be justifiable. The only word of caution I have is this: what will the effect on research be if authorities start making grants contingent upon obtaining consent from all subjects?

PART SEVEN

PATENTS AND
COMMERCIALISATION IN AN
e-RESEARCH WORLD:
NEW MODELS

SCIENCE AS SOCIAL ENTERPRISE: THE CAMBIA BIOS INITIATIVE

Richard Jefferson¹

Nearly four billion people live on daily incomes lower than the price of a latté at Starbucks. Most of them make dramatically less than that—and from that income, they must acquire their food, their medicine, their shelter and clothing, their education, and their recreation, and they must build their future and their dreams. Their lives, and the quality of their lives, hinge on biological innovation.

Biological innovation is the ability to harness living systems for our social, environmental and economic well-being. It is the oldest and most fundamental form of human innovation, involving as it does the getting of food, the striving for health, the making of homes, and the building of communities. The wealth created over the millennia through the domestication and husbandry of plants and animals has powered human society.

Of all areas of biological innovation, agriculture is the most important, affecting our environment, our health, our economies, and the fabric of our societies. The world's poorest nations depend largely on agriculture for their economic survival as well as their food, fuel and fibre. The challenges of innovation to create and sustain productive and environmentally sound agriculture are even more pronounced in these societies. Any failure to do so has enormous implications for the global community, over and above the social, economic, and environmental impacts.

For thousands of years biological innovation has been informed and guided by keen observation and the accumulation and sharing of generations of empirical knowledge. Farmers selected better crop varieties and livestock breeds, and developed management strategies to

¹ Richard Jefferson is the founder and CEO of CAMBIA-BiOS. This chapter was first published as an article in (2006) 1 (4) *Innovations: Technology, Governance, Globalization* 13–44.

maximise their performance. Seeds were shared as a practical matter of survival and each improvement formed the basis for further innovation. Because seeds of most crop plants breed true, the ease of sharing, and the barriers to doing so were minimal. As with digital information, it is hard *not* to share, and hard to impose limits on sharing, so norms evolve to maximise value from this inevitability.

But the post-Enlightenment explosion of possibility that began when the unprecedented power of science became focused on food, agriculture, health, medicine and environment seemed to dwarf all previous attainments. And indeed in the past hundred years, with the advent of genetics, the pace has been gathering; the last 30 years has seen an unprecedented dynamism in life sciences that is being hailed as a 'biotechnology revolution'. But in this revolution, biotechnology is rarely being applied to the critical issues of alleviating poverty, eliminating hunger, stewarding natural resources, and preventing or curing the diseases of the disadvantaged. The margins are small, the markets are modest, and the challenges are great. Are the paradigms and practices that have emerged to harness science for society sufficient to engage, and even solve, these seemingly intractable problems?

Today control over agricultural biotechnology is effectively limited to a few multinational corporations who integrate seeds, agrichemicals, and biotechnology. This disturbing consolidation of power is matched with a trend toward 'me-too', big-ticket 'innovations' of remarkable dullness. How many herbicide-tolerant big acreage crops are enough? Similar trends are surfacing among the large pharmaceutical companies, collectively known as 'big-pharma': how many blockbuster lifestyle drugs does society need?

Within the value system they respect, and according to their own success metrics of profitability, big agriculture and big pharma are not abject failures, but they surely are not enough.

To address the myriad challenges of agriculture, environment and health that are local in nature and modest in market or profit margins will require vigorous, competitive, local-scale small to medium enterprises creating a business and innovation ecology. It will also require a biological innovation culture where the costs of innovation are low, and the power and relevance of technology are high. It will require leveraging the contributions of diverse people and institutions to create

tools that better engage science into an integrated and economically sustainable social agenda.

The mission of CAMBIA, of which I am the founder, is to advance this set of required capabilities so that biological innovation can address the human challenges of the 21st century; the BIOS (Biological Open Source) Initiative is CAMBIA's mechanism for achieving its mission.

The term 'open source' describes a paradigm for software development associated with a set of innovation practices. The concept evolved out of the 'free soft-ware' movement, and is often merged into the expression 'free and open source software' (see text box.) Several features together qualify a project as 'open source'.² These include full disclosure of enabling information including documented source code and the use of legal instruments such as copyright licenses to confer both permissive rights and responsibilities; they bind contributions into a commons that is accessible to all who agree to share alike. Typically, certain practices and cultural norms are associated with distributive innovation, although this is by no means required; some very successful free and open source software projects have only a few serious contributors, while others have thousands.

Extraordinary efficiencies occur when the tools of innovation are shared, are dynamically enhanced, have increased levels of confidence (legal and otherwise) associated with their use, and are low or no-cost. Rent extraction from the process of innovation is reduced, transactions costs are minimised and developers focus their resources on creating revenue by providing products and services and enlarging markets.

² For example <www.opensource.org>.

How Do You Make Money in Open Source?

Free and open source software has rapidly engendered highly productive and profitable business models that create value from the non-rivalrous³ use of software components. Examples of such software include the famous Linux operating system, the Apache web server, databases such as MySQL, myriad programming languages such as Perl and Python, and the Firefox web browser. These types of open source projects, co-developed by thousands of programmers, and shared through creative licensing which demands covenants of behaviour rather than financial consideration from the licit community of users, have transformed the information and communications technology (ICT) sector.

Most of the high-profile free and open source software projects that have affected both the sector and the public's imagination have been 'tools' and platforms, rather than end-user applications. These allow users to build fully commercial web applications, with high functionality, on robust, dynamic platforms, with no reach-through financial obligations. The economic success stories of free and open source software thus are not Linux and Apache, but eBay and Google. The business models that are shaking the ICT world are not the modest ones selling support for open source products, such as Red Hat Linux. The signal successes are commercial enterprises that create wealth by providing new social value. Many ask, 'How do you make money in open source?' The answer: you make money not by selling open source, but by *using* open source.

This concept is fully generalisable—although clearly the specifics are not—and a large part of CAMBIA's BiOS initiative explores and extends the software metaphor. BiOS strives to create new norms and practices for dynamically designing and creating the tools of biological innovation, with binding covenants to protect and preserve their usefulness, while allowing diverse business models for wealth creation, using these tools.

³ In economics, a good is considered either rivalrous (rival) or nonrival. Rival goods are goods whose consumption by one consumer prevents simultaneous consumption by other consumers. In contrast, nonrival goods may be consumed by one consumer without preventing simultaneous consumption by others. Most examples of nonrival goods are intangible goods (Wikipedia, 2007).

In the first part of this chapter I discuss the simultaneous burst of knowledge in molecular biology and the precipitous decline of a commons of tools, using examples from plant biotechnology. I develop a practical model of innovation, highlighting how biological innovation is stymied or deflected to high margin applications if tools are not freely available, continuously improving and embodying the permission to deliver work product into markets. I explore parallels, divergences and resonance with open source paradigms in software engineering. The rest of the chapter focuses on CAMBIA BiOS Initiative activities: the BiOS Framework, the Patent Lens, and the BioForge, and the creation of a ‘commons of capability’ through which new actors, including farmers and small-to-medium enterprise, can use science to create viable innovations relevant to their needs.

POWER, TOOLS AND THE COMMONS OF CAPABILITY

Twenty-eight years ago, I began a project to develop a set of tools—of techniques—in molecular biology that could help researchers in that field visualise how genes and cells functioned. Like virtually all scientific work, and most technology development, it was inspired and informed by what came before. And like all tools and methods, it depends on the use of other tools and methods.

Some years earlier, Ethan Signer, Jonathan Beckwith, and others had made a remarkable contribution to our toolkit for understanding how genes worked in bacteria. They conceived of a single tool that would allow scientists to learn how genes turn on and off in a bacterium. The tool ‘hooked up’ the beta-galactosidase gene (called *lac*) for which they had simple measurement tools and assays, to another gene (called *trp*) for which measurement was difficult, but whose behaviour they were keen to understand. In so doing, they measured the *trp* gene by actually measuring *lac*. This *tour de force* of microbial genetics used publicly available technologies and methods—in fact it was then unthinkable that there would be any other kind. This occurred well before the advent of recombinant DNA, which now allows apparently sophisticated genetic experiments to be done very simply. And it occurred well before anyone had even contemplated patents on life sciences.

Years later, I thought, why not use the same concept to understand how cells in animals and plants work? Why not have the organisms talk to us about their environment, through their genes? I set out to develop a parallel system, using a different enzyme and gene that could function in these new organisms. The one I chose was prosaically called GUS.

As I worked, I became increasingly aware that the availability of tools, and their capabilities, completely dictated the science that was done, and who was doing it. As an undergraduate at the University of California and the University of Edinburgh, I worked in some of the key laboratories responsible for inventing recombinant DNA methodology. I watched, time and again, how an entire field of scientific endeavour would almost instantly change course when a new technique—tool—was provided.

When I first developed the GUS technology, the scientific community I was originally working within—which studied animal embryo development—was not very interested; the tool just wasn't needed much. My first paper on this topic was received with an ill-stifled yawn. But I moved my interests to plants and agriculture, during the heady dawn of plant molecular genetics.

Efforts to transfer beneficial genes into key crops such as cotton, soybean, maize, and rice were running into a brick wall. There was no way to visualise success, nor to measure and improve on first steps. The GUS reporter system made visualising genes and their action in plants very easy and efficient—it was proving to be a very powerful tool at the right time.

In 1985 I arrived for my postdoctoral research at the Plant Breeding Institute (PBI) in Cambridge, England, a vigorous international group of colleagues who were at the cutting edge of technology development and exploration in molecular plant sciences. The Plant Breeding Institute was also one of the few sites in the world that combined the patient and disciplined craft of successful agricultural innovation, such as plant breeding and agronomy, with the impatient and fermenting world of molecular biology. As well, the Plant Breeding Institute was still at that time an entity focused on the public good, a non-profit institute that earned substantial income for the U.K. government through royalties on its own crop varieties.

At Plant Breeding Institute, my colleagues⁴ and I designed and conducted the first field test of a transgenic food crop. It was also the first BioSentinel experiment: a gene we wished to study was fused to the GUS gene, to conduct a field trial asking a fundamental question about how genes act under field conditions. We used public money, in the public sector, to ask a fundamental question for the public. The field was planted on 1 June 1987—completely by chance one day before Monsanto's first field trial. The lessons of the field trial were fascinating. We found that gene activity in a field is extraordinarily variable, and our preconceived laboratory-based notions of how genes worked would turn out to be very inadequate when dealing with field populations. Our technology, though cutting-edge, was not up to the questions that real-world agriculture presents.

The Plant Breeding Institute was an international institute, with students and scientists from all over the world. The institute had a reputation for brilliant wheat breeding and genetics, so most of the countries whose agriculture depended on cereal production would send their scientists to us for training. Many of the students and postdoctoral fellows in the Molecular Genetics department came from India, Pakistan, Turkey, the Middle East, China, Africa, Latin America, and Eastern Europe. Most of them indicated that this period in Cambridge was their one shot at career establishment. If they published a paper or two in a good journal, they had a reasonable chance of employment back home. And some of them confessed that they likely would not be able to use the new biotechnologies to effect any change in their home agriculture or economy. Not only did they lack the finances and infrastructure to make use of these high-tech tools, but the tools were better for science than for problem solving.

These people were exemplary of perhaps the most crucial but neglected resource for social advancement through science: dedicated and capable

⁴ Mike Bevan, my principal collaborator, went on to play a key role in coordinating the public sector sequencing of the *Arabidopsis* genome. *Arabidopsis* is the workhorse model plant of biotechnology, and was the first plant to have its entire DNA sequence described in the literature. The public efforts to create a public good, like some of mine, were likely co-opted by the secretive wholesale filing of patents on the *Arabidopsis* genome by Mendel Biotechnology, an affiliate of Monsanto. These patents have only recently surfaced <www.patentlens.net> but pre-dated the public effort by as much as two years, thus potentially capturing or hijacking much publicly-funded work, through a legal, though unpalatable practice called 'after-claiming'.

people. I observed, however, that instead of using their own experience to inform the science that was being done and the technologies being developed, their own world-views and self-images were rapidly aligning to the incentive and reward system of the prevailing and fashionable science trends. And their energy to change the options in their home countries was dissipating.

By early 1987, after intensive experimentation in-house, we had assembled hundreds of copies of a GUS kit of dozens of DNA molecules and a comprehensive ‘how-to’ manual. I rewrote the big ‘GUS Manual’ and sent it to a mass-mailed newsletter called *Plant Molecular Biology Reporter*, which was distributed free to thousands of scientists rather than initially publishing a peer-reviewed scientific paper, which I eventually did.⁵ The grapevine is also a powerful communications tool in science; soon many people were hearing about this new technology that would let them see the cells and tissues where their gene was functioning. It would also allow let them optimise gene delivery experiments; this was an urgent priority for both industry and academia. At that time no important commercial crop had been genetically engineered, so requests started flooding in for the GUS system. And I started sending out hundreds, even thousands of samples, and the User’s Manual, all with no licenses, to scientists in dozens of countries, in both the private and public sectors. I only included a letter saying that while I had filed for a patent on the system, I wanted everyone to use it, and royalties—if any resulted—would go back to creating the next generation of technology.

I sent the kit to scientists at Agracetus in Wisconsin who were working, with little success, on transferring genes to soybeans. They had no idea if the genes they were introducing with their new process were actually making it into the right cells. One of those scientists, Paul Christou, told me of their thrill when they were able to immediately visualise gene transfer with the blue colour of the GUS test, and soon succeeded at introducing genes into soybeans for the first time. And they could only

⁵ R A Jefferson, T A Kavanagh and M W Bevan, ‘GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants’ (1987) 6(13) *European Molecular Biology Organization Journal* 3901–7. Apparently it has been read often, as it has been cited in the scientific literature thousands of times. To our delight, however, the user’s manual in *Plant Molecular Biology Reporter* has been similarly cited, and likely more influential, in the precursor to the Open Access publishing movement.

do it with GUS, which also had no apparent restrictions. They were delighted, of course, as was Monsanto, for whom they worked.⁶

That work with GUS turned out to be the single biggest money maker in plant biotechnology, possibly ever in agriculture. Monsanto developed its RoundUp Ready™ soybean line, which it ultimately used to breed most of the transgenic soybean plants now covering the world, using GUS to select plants.

Within a year after we began widely distributing the GUS technology, hundreds of new avenues of plant science were emerging. Within two years, breakthroughs in maize, soybean, cotton, and many other crops occurred. New technologies were invented that used the tool in its very creation and optimisation, such as particle bombardment (the tool that Agracetus had been exploring) and critical improvements were made to core technologies such as gene transfer by *Agrobacterium*. GUS demonstrated that one powerful new tool, widely distributed, could rapidly change an entire field.

The idea of *intentionally* changing the directions of inquiry and the demographics and economics of problem-solving by designing and providing new tools would shape the next 30 years of my professional life. With increasing exposure to the realities of practical agriculture, intellectual property, policy and business, my definition of ‘tool’ matured. It came to include not just the technologies needed for scientific investigation, but also the critical normative, economic, policy, legal and business instruments to convert investigation into socially and economically sound innovations. A business model really can be a tool.

Enclosing the Toolkit: The Case of *Agrobacterium*

But while this period hinted at the vast potential for new tools emerging from molecular biology to lead to rapid innovation, it also saw the rush to privatise the kinds of tools that had always been seen as a commons, as exemplified by the adventures of *Agrobacterium*. When I started to work at Plant Breeding Institute, plant molecular genetics was in its infancy, and only three or four major institutions had serious capability

⁶ Monsanto later engaged Agracetus in a heated patent battle for the right to do genetic manipulations in soybeans, and ultimately purchased Agracetus and its patents. At this point the patents owned by Agracetus ceased being seen as reprehensible and unfair, and were defended as pillars of rectitude.

in this nascent field. All of them were using *Agrobacterium*-mediated transformation as their fundamental tool for transferring genes to plants. Several years earlier, several public research teams had discovered an astonishing biological phenomenon.⁷ A soil bacterium long known to be the agent of a familiar plant disease called crown gall was found to cause these tumours on plants by a hitherto unforeseen mechanism. The bacterium—*Agrobacterium*—actually inserted into the plant, by ‘natural’ genetic engineering, a component of its own genome, and in so doing reprogrammed the plant to produce a ‘gall’ and new food for the bacterium. This phenomenon, a sort of biological Trojan horse, was thought to be unique in the biological realm. And everyone in plant biology saw that it was to be a critical tool in the development of new options of biotechnology.

The groups that first made the discoveries were all in the public sector, funded largely by public monies; they could all see that *Agrobacterium* would be a fundamental tool of the field. In spite, or perhaps because of all this, the gold rush for patenting started. And not only did the pioneer groups in the field file patents; over the next 20 years over a thousand patents were filed—and granted in many nations—that covered various aspects of *Agrobacterium*-mediated gene transfer. Some were so minor and trite as to be laughable were they not presumed valid by law, but they still produced a thicket of rights, nearly impenetrable even to the specialist.

And of course the pioneering patents were fought over viciously. To monetise the patents, the rights were sold to the highest bidder. But the rights were not clear; bitter wrangling over primacy with the fundamental patents continued for almost 20 years before any legal clarity emerged. Of course the winning bidders ended up being large multinational companies, notably Monsanto (either directly or by acquisition); and in most cases the payments to universities and institutes were negligible or even negative. But the effect of increasingly consolidating these patents in a few hands was anything but negligible.

Soon, public and private sector scientists were patenting their developments as a matter of course. Some of these findings became

⁷ These scientists included groups led by Mary Dell Chilton, Marc van Montagu, Eugene Nester, Jeff Schell, Pat Zambryski and others, at the University of Washington, the University of Ghent, the Max Planck Institute, and elsewhere.

powerful patent estates that potentially blocked most of the world's agricultural enterprises from using these tools without permission, often at any price. For example, Japan Tobacco discovered and patented a method to use *Agrobacterium* to transfer genes into rice and other cereal crops.

The case of *Agrobacterium* was repeated with many subsequent technologies, ranging from genetic selections, to the wholesale patenting of promoters and genes,⁸ to gene inactivation technologies (such as RNAi and co-suppression). Again, the contents of many patents were breathtakingly obvious to all practitioners in the field, but for small to medium-sized enterprises these patents still served as a real disincentive to innovate. They also extracted huge rents from industry, and raised transaction costs to an unbearable level, mostly because the patent landscapes were so opaque and complex. This trend has accelerated markedly and now applies to medical as well as agricultural technologies. The consequences are clearly that only the biggest-ticket targets are getting attention. But blockbusters alone don't make for good agriculture, good environmental management or good public health.

In 1985 the sector was viewed as exhilarating, entrepreneurial and vibrant, with almost unlimited possibility for doing good in world agriculture; within a decade or so it had all but stalled into a corporate oligopoly, with vertical integration, ossified and oppressive business models, and massive patent portfolios tying up almost every key technology and platform used in the field. And though nearly all the pioneering discoveries were made in the public sector, they were not reserved for public use or for the small-to-medium enterprise sector that the public trusts. It is no surprise then that the public now views the entire agricultural biotechnology sector—as manifest in the outcry against GMOs—as being a tawdry exercise in failed promises, industry consolidation, public sector abandonment and simplistic agendas. Perhaps the greatest crisis that has emerged from this corporate control of problem-solving in agriculture is that the public now seems to have very little confidence in the use of *any* science in agriculture! This has indeed been a case of throwing the baby out with the bathwater.⁹

⁸ See forthcoming 'Patent Landscape on Patent Genomes'.

⁹ R A Jefferson, 'Transcending Transgenics: Is there a baby in that bathwater, or is it a dorsal fin?' in Phil Phardey (ed), *The Future of Food* (2001) 75.

Biotech Bazaar: Tools for Sale

At the Plant Breeding Institute, I was working with colleagues from scientific cultures that had historically used the discoveries and technologies that came before to grapple with the next generation of scientific challenges, with the tacit understanding that this process would naturally yield real-world solutions, such as plant varieties and agronomic processes. After all, the Plant Breeding Institute paid its way in the world by doing just this.

But that world was collapsing. The distinction between discovery and invention was being blurred as patents were filed on each component; that process entirely altered the dynamic of translation into true innovation: delivering the products of science and technology to the marketplace. It was now possible to control the tools and platform discoveries themselves, not just the products that they created.

In the early 1980s with the passing of the *Bayh-Dole Act*, universities in the United States were actively encouraged to patent their work products. The *Act's* fundamental policy goal was to see publicly-funded science and technology better used by society, by encouraging industry to adopt it. The trend of public agencies using the patent system exploded internationally into a filing frenzy. No one foresaw then that the fragmentation of the platforms and tools would make it so complex, so expensive and so intractable to assemble the 'freedom to operate and freedom to innovate'. Nor did we see that the resulting innovations themselves would be so few, so stodgy, and so slow to reach the marketplace.

At almost the same time, the advent of recombinant DNA and the ability to determine DNA and protein sequences massively increased scientists' ability to explore, understand, and manipulate living systems, or at least living organisms. So every new life sciences discovery could be, and often was, dressed up as an invention and subject to patent; as the patent claims were granted, they cast a huge net over the possible options. Public sector coalitions would frequently compete with private big-science, and who usually won the plum of patent monopoly? The privatised efforts. Was this right, or necessary?

I began my own foray into patents and their importance when I arrived in Cambridge in 1986. I discovered close relationships between some large companies and the public-sector institute where I was based,

shaped by personal histories and friendships. I didn't view this as a bad thing. I shared all my ideas and technologies with them from the outset. In fact, I shared with pretty much anyone who was interested, thinking that—in economic terms—my ideas were non-rival; sharing didn't cost me the ability to use them myself. How wrong I would later prove to be.¹⁰ And how times were changing.

One company, ICI,¹¹ was keen to use GUS in its commercial development programs; like many companies it was mostly interested in having clear rights to do so. ICI suggested that I patent my technology so it could be sure it would have access to GUS in the future. I didn't understand the logic at the time, but I took the first steps and filed a patent in the United Kingdom and the United States, with a filing date in 1986. The University of Colorado, where the first stages of the work had been done, had waived its interest in patenting it.

Thus began a long and painful learning process of partnerships with powerful attorneys in which I watched patent-craft by The Masters. It took almost seven years for my first patent to issue in the USA, and nine years for the one with most of the valuable claims. Even when it was issued, complex agendas and issues¹² kept me from licensing the patents or even having a clear title for quite some time. This delay wrought havoc with my ambitions to use patents to create and fund CAMBIA, and when revenue did come in, it was in sporadic bursts, and barely in time to make payroll.

As a technology, GUS has had a surprisingly long shelf-life, and is unusual in being a largely stand-alone technology. If one has the 'right' to put a gene into a plant, GUS remained a useful and legally usable tool to monitor that gene and its activity. But it turned out that even that right, the legal permission to transfer a gene to a plant, proved to be a critical and contentious issue because patents are opaque and licensing rights even more so, and because advances in the life sciences are so interdependent.

¹⁰ See appendix on positive selection.

¹¹ Imperial Chemical Industries; its plant work was later absorbed into Zeneca and then into Syngenta.

¹² More details on the complexities of this period can be found in Richard Poynder, Interview with Richard Jefferson (online interview) *The Basement Interview: Biological Open Source*, <<http://poynder.blogspot.com/2006/09/interview-with-richard-jefferson.html>>.

Wheels and Spokes: The Interdependency of Technologies

The patent system is so complex it is almost awe-inspiring. Single patent documents can run to hundreds of pages, with arcane language that few understand, and rights that courts interpret and re-interpret on the fly. Thousands of these can exist in a single field of innovation, with many thousands more latent in the system. One or two—or none—may be, or may unexpectedly become, dominant. Fundamental biological processes, such as the ubiquitous gene-regulation mechanism, RNAi, have been patented. Most of the important genes of many important organisms—humans, rice, maize, mice—have been subject to patent applications and sometimes grants, many of them contestable by many separate claimants. The platforms on which we must build are privatised and enclosed, but the owners and their ambitions are completely unclear; the platform for future innovation is built on shifting sand.

But worse, while the ownership of the ‘patent’ itself is usually a matter of public record, the ownership of the *rights*—the most important feature of a patent—is completely obscured. Nowhere, in most jurisdictions, is there recorded or available the patterns of power: who owns what rights. A university may own hundreds of patents, and may have sold off the rights to any of the useful ones, but who bought them? The answer is rarely clear.

When a small company licenses a patent, or develops its own patent portfolio, to whom has it licensed and on what terms? The patterns of power and ownership are as important—and in the aggregate perhaps more important—than any other feature of a patent grant. And yet we have no public information whatsoever, except in piecemeal and scattered disclosures. Some jurisdictions, including Brazil and France, do impose a responsibility on licensees to disclose—at least to the patent office. But most do not. And none make it easy to find this information. This makes it difficult, if not impossible, for a researcher in a small- or medium-scale enterprise to assemble all the licenses or capabilities needed to refine and adapt a tool and ultimately to create an innovation that will help meet basic needs.

And researchers need this information because few discoveries stand on their own, and even fewer inventions. Not only do they each depend on the pre-existing knowledge base; they almost always incorporate components of many other technologies in their execution. This is

particularly true of ‘meta-technologies’, tools and technologies with broad effects used by communities of innovators quite distant from the tool’s original inventor.

Consider the wheel, perhaps a six-spoked wheel. In some ways, it is the most fundamental and important tool in society. It has countless uses unanticipated by its inventors; most were made by people who are not wheel-builders. The wheel is only useful when it is used for something, such as moving a cart; its economic value to society lies not in the price of the wheel, but in the wealth created through the use of the wheel.

If it takes all six spokes for this wheel to turn, and each of these spokes is potentially different in some way, we have a good metaphor for a modern biological technology. Increasingly, biological technologies are not self-contained; rather they are rather interdependent technologies that require multiple key methods and components to function. If one spoke is withheld, no wheel is built. If one spoke is broken the wheel will jam. And then the cart cannot move forward. By analogy, the most powerful technologies can be considered as ‘wheels’, requiring a number of ‘spokes’ to function. For instance, the ability to transfer a gene to a crop plant may require dozens of individually protected, discrete technologies. Denial of access to any one of these ‘spokes’ obstructs not only the use of the technology, but its improvement. Only when the core technology is in place, with full functionality, can it be subject to iterative and cooperative shaping to meet diverse users’ needs.

Unfortunately, even placing one or more key methods or components into the public domain allows no leverage to bring other components into a collective whole with broad access. Virtually all the practices of academic scientists promote the belief that ‘good science’ can, almost by magic, transform itself into public or private goods. It can’t. In fact, by failing to deliver such goods with broad and preserved access, the public sector science community is complicit by neglect, because the true stranglehold rests where much less public sector effort is expended: in the process of converting invention and discovery into innovation, by building and using wheels.

But we can change this landscape, if we provide one or more of the spokes to all the wheel-builders and users with covenants of behaviour, rather than financial consideration (outlined later as BIOS licenses). If a

user can access a spoke only by promising to share spokes, or improvements, then the whole logic can change.

This is where we find the leverage: change the logic of copyright licenses in software to allow free and open source software to exist, and do the same for patent licenses or Materials Transfer Agreements (MTAs) in BiOS. Then we can regain a full complement of spokes, and see the ‘wheels’ of real innovation turn rapidly and deploy on many roads, creating wealth through their use.

How Fear, Uncertainty, and Doubt Can Deter Innovation

Uncertainties over intellectual property rights undermine the long-term and sustainable pursuit of innovation by making projects look more risky to potential partners and investors. This risk combines with others characteristic of early stage technology development: lack of a fully-specified business model, concerns over potential technology effectiveness, and the absence of a well-established delivery channel. Together they generate the fear, uncertainty and doubt (FUD, in the awkward but widely used acronym) that is the core impediment to technology development. Currently, every worldwide industry that depends intensively on science and technology experiences FUD. Sometimes a competitor is the focus; sometimes the bleak patent situation alone can lead an investor, client, customer and/or the public to lose confidence in the prospects of creating a viable technology-driven enterprise.

In the face of the uncertainties associated with the complex and opaque patent situation, multinational private-sector firms have responded by acquiring large IP portfolios and negotiating cross-licensing arrangements to obtain platforms of enabling technologies. Even so, these companies still often find themselves with constrained freedom to operate. Faced with the uncertainty of patent rights, they seem to be involved in a sort of mutually assured destruction.

In contrast, the public-good sector, and small-to-medium enterprises have only fragmentary portfolios, often made up of publicly-developed technology and modest non-fixed capital pools that they believe can be expanded by their eagerness to license them out, but they are at a grave disadvantage; they face a monopsony.

Unfortunately, this approach not only destroys public value and confidence; it is also ineffective in ensuring a sustainable private competitive advantage. As the expense of sequestering intellectual property outside the public domain in iterative patents has increased, some leading technology firms have decided that an open source model may yield higher private, as well as public, returns. A notable example is IBM Corporation; in a bold recent move it is stimulating a universally accessible ‘protected commons’ of patents in a pool available for any open source development. As the world’s largest patent holder, IBM could be viewed as a ‘rights maximalist’; over 500 of its key software patents have been made available to all—including competitors—who choose to use them under open source rules. Within days, Sun Microsystems followed suit with another 1600 patents, and a myriad of other companies are doing the same. The snowball effect continues, as companies realise that their sector makes progress when the standards and the toolkits are clear, open, of high quality and consistently available.

Clearly, true wealth creation will come not through extracting rent from a tool, but through *using* a continuously improving toolkit, with continuously decreasing costs of innovation and a continuously expanding group of tool users. Diverse and prosperous agriculture, robust public health and sustainable natural resource management are the publicly valuable goals we must keep in clear sight. The tools associated with their improvements must be plentiful, powerful and affordable.

As the ICT sector realised, we also need an open source movement in biological innovation that can empower public and private sector innovators with the tools, platforms and paradigms to allow rapid and efficient life-sciences innovations for neglected priorities and new opportunities.

CREATING CAMBIA, MAKING CHANGE

In the mid-1980s, when I first formulated the ideas that became CAMBIA, I did not intend to build an institution; I spent much time between 1987 and 1990 trying unsuccessfully to convince universities or

later the United Nations or the CGIAR¹³ system to take on and host CAMBIA's mission. But the complexity and edgy nature of the mission, the need to integrate diverse skills and strategies, and the entrepreneurial spirit, ultimately required an independent base.

In early 1992 I moved to Canberra, Australia, to begin a project on behalf of the Rockefeller Foundation, troubleshooting its rice biotechnology network in Asia. At this point CAMBIA was not a legally incorporated body, but had reams of letterhead and surprising credibility. Our job was to travel to virtually every laboratory in the developing world that had Rockefeller Foundation support—and over the next eight years this must have been hundreds—to help develop, improve, and apply their biotechnology capabilities, especially as they pertained to rice molecular biology. We developed and provided to many hundreds of labs—perhaps over a thousand—the most effective and widely used ‘vectors’ in plant molecular biology, the pCAMBIA series, and provided courses and workshops in the science and increasingly over time, in intellectual property management. In hundreds of working visits to China, Indonesia, India, the Philippines, Thailand, Vietnam and many other countries of Asia, Africa and Latin America, we forged a sense of the possibilities if we had new types of technologies, and new communities to improve and share them.

During these years, as we became more sophisticated about licensing and understanding the patent systems, we also became more aware of the yawning gulf between biotechnology rhetoric and innovation realities in most of the world. On the one hand we saw a large, untapped population of dedicated and knowledgeable problem solvers, committed to solve problems of real substance to their countries and peers—but they lacked the usable technologies that would improve their situation. We also saw that the science itself was not up to the job: the research being conducted in the early days of plant molecular biology (and sadly still now) is intensely reductionist, whereas the problems of agriculture and society are integrated into complex systems. On the other hand, if we could design and provide tools that fit the problem and the hand of the tool-user, we could rapidly and effectively change the entire platform

¹³ The Consultative Group on International Agricultural Research, <www.cigar.org>, a consortium of 15 agricultural research institutes and many governments, is the principal non-profit entity engaged in agricultural development through science for poverty reduction.

of problem solving, as long as the tools were dynamic and could embody the permissions to integrate into real-world innovation. CAMBIA was conceived to integrate and to address these issues.

Outlined in the earliest CAMBIA prospectus was the premise of using patent revenues to create a sustainable funding base. We surmised that we would ask a fair, tiered licensing fee of each company that was using the technology, proportionate to their ability to pay. A big company pays a lot, mom-and-pop companies pay peanuts, developing countries pay nothing. Then we would use the resulting revenue stream to invent and distribute the next generation of technology. At the time it looked like a logical and efficient way to move the sector forward with fair and open competition, not for the capability to innovate, but for the innovations themselves.

This worked to some extent, in that CAMBIA exists and might not have done so without patent revenues. Companies that licensed the technology range from giants like Monsanto, DuPont, Pioneer, Bayer, BASF, and Syngenta down to entities as small as the Hawaiian Papaya Growers Cooperative. But we also realised we could not generalise or scale it as a business model in the current climate of fragmented rights and capabilities. The transaction costs of negotiating licenses, as more and more ‘spokes’ were required to move forward, would simply be impossible to bear for any but the highest-margin applications.

CAMBIA addresses these challenges through three interdependent activities:

1. *The BiOS Framework* creates, validates and promulgates licensing tools, along with the norms and new business models to make use of strategies for ‘open source’ creation, improvement, and sharing of enabling technology.
2. *The Patent Lens* is a platform to focus, understand, and investigate the patent rights and to inform practitioners and policy-makers.
3. *CAMBIA’s own research* into creating and distributing key ‘pump-priming’ enabling technologies is made available through our online interface, the BioForge.

The BIOS Framework

Biological Open Source is a nascent movement, evocative of the transformative changes in information and communications technologies (ICTs) wrought by free and open source software (FOSS). The two movements share some goals: seeing transformational effects on a sector, and increasing the democratic involvement in problem solving; we are learning many lessons from the software world, and will continue to. But it would be a mistake to push the comparison too far. BIOS concepts have emerged from 20 years within the life sciences and human development culture, to address the needs and challenges of biological innovation.

The idea of using patent licenses not to extract a financial return from a user of a technology, but rather to impose a covenant of behaviour, is the single feature of BIOS that is most resonant with Free and Open Source Software. We¹⁴ worked with small companies, university offices of technology transfer, attorneys and large multinational corporations to understand their concerns and experiences, and then create a platform to share productive and sustainable technology.

Patent Lens: A Platform for Understanding IP Landscapes

CAMBIA's Patent Lens includes one of the world's most comprehensive full-text searchable databases of patents; cost-free and available to anyone, it has a seven-year history of continued growth in features and power. It incorporates the full text of applications and granted patents from the U.S. Patent and Trademark Office, Patent Cooperation Treaty (PCV) database, European and Australian jurisdictions, and their status and family relationships in many dozens of countries. Its fast and user-friendly search engine has a nuanced interface and presents common and harmonised data structures so that these jurisdictions can be searched simultaneously. The Patent Lens is becoming an increasingly important resource as the fee-requiring 'value-added' patent data providers continue to consolidate.

¹⁴ Dr Marie Connett, CAMBIA's Deputy CEO, a scientist, patent agent, and IP Manager, jumped into the deep end when she joined in 2005, and found herself working round the clock on creating the license, consulting with dozens of technology transfer professionals, lawyers, industry colleagues and scientists.

Because no national patent office has taken on the task of harmonising collections over many jurisdictions, the role of the ‘patent clergy’ remains central, and the gate-keeper functions of the information providers remain onerous. National and regional patent offices provide quite variable free patent searching; some are appallingly primitive while others, like the European Patent Office, are quite sophisticated. Patent offices, however, have complex relationships with commercial providers such as Thomson, which actually provide the patent offices with integrated searching functions for their own in-house use. To further complicate the situation, commercial providers have been calling for a reduction in the role of national patent offices as ‘value added’ providers. The need for a public good provider has never been greater.

Patent Lens focuses on user-adaptability, integration, annotation capability and availability to the world community for free; these key features render it particularly helpful in efforts to restore public good and transparency as the *raison d’être* of intellectual property systems.

Technology Intellectual Property (IP) Landscapes

IP Landscapes are analyses of key platform technologies, and the IP positions associated with their development and use. They build on and use the patent database, but include much more than a collection of relevant patents. Each landscape is a searching and analysis effort involving many person-months, by CAMBIA staff and soon others, who have particular knowledge of the science and technology and of patent claims. Typically, patent ‘professionals’ within law firms accumulate billable hours by providing the same information over and over for different customers, and charging full fees again to update them periodically. Increasingly we wish to do something no fee-requiring patent data provider will ever do: turn the landscapes into living repositories of constantly updated information, so no more updates will ever be required.

The goal is to use the harmonised datasets to create a facility where distributed and diverse users can generate, link, and dynamically annotate patent landscape analyses through web interfaces. The landscapes will ultimately become maps and decision support tools so users can distinguish green-fields from minefields in the long path from discovery to practical delivery of an innovation.

We have created a substantial number of such landscapes, in an early, hypertext-linked but basically flat structure. But we aim to enable the preparation of many more, by many people, by leveraging informatics to create ready frameworks and linkages between world patent literature and such resources as PubMed Central, and Google Scholar whose relevance engines can enrich the process. Ultimately we see the navigation of technology landscapes as being a critical feature in research and development decision making, but people will only use them when their costs, in both time and money, are negligible and the relevance and utility of the guided decisions are clear.

Patents, Policies & Practices

This component includes tutorials that guide users in reading and interpreting patents; the aim is to make novices more sophisticated about the nuanced realities of intellectual property, particularly patents. It also includes Policy & Practices papers that describe and advocate for informed and productive changes in international, regional and national forums and laws.

The goal is to forge a learning resource that participants in innovation systems at all levels—scientists and engineers, business and legal professionals, citizens and policy-makers—can use to learn of critical and timely issues relevant to improving the public good and social and economic value by engaging with the patent system.

The standards of modern patents are widely viewed as execrable; though many patents are presumed valid by law, they are at best frivolous and often egregious. We aspire to provide the public with tools to recognise and overturn such patents where they undermine progress or are being used without a long-term and well-articulated stake in industry or society.

The basic premise underlying that license is that we would not charge any fee for use of the ‘basket’ of technologies with the patent estate being offered. By making the license cost-free, we hoped to induce the most valuable contribution to the license community: ‘freedom to innovate’. In exchange for full, unfettered commercial rights to our technologies, licensees are required to comply with three conditions:

- They will share with all BiOS licensees any improvements to the core technologies as defined, for which they seek any IP protection.
- They agree not to assert over other BiOS licensees their own or third-party rights that might dominate the defined technologies.
- They agree to share with the public any and all information about the biosafety of the defined technologies.

Several further features of BiOS Certified licenses are very important:

- The definitions are critical. The core capabilities (enabling technologies, platforms) and their scope must be carefully defined to allow confidence in the development of viable business models that use these BiOS licensed technologies.
- The BiOS License structure must be scalable, and it should be generalisable, capable of development within these guidelines, and overseen by diverse institutions. We recognised that different technology sets have very different implications in the innovation chain, and that the agreement must accommodate different sectors (for example, agricultural and medical) and different economic circumstances (industrialised and less-developed countries). Therefore we developed a suite of licenses around several different enabling technologies CAMBIA developed. We created them around our own technologies to have first-hand learning platforms from which we could generalise and help others create their own BiOS-Certified programs.

As we have gained experience with our first-generation licenses through the concerns and suggestions of many licensees and potential licensees, we have aimed to create a 'brand' of Biological Open Source (BiOS) that is independent of institution. The BiOS certification program will help ensure that core BiOS characteristics are sculpted into forms that allow institutions to preserve their own cultures and priorities. They may do this through the medium of patent licensing or through materials transfer agreements (MTAs), a common form of bailment used to

provide materials for life sciences research, such as bacterial strains, plant lines, cell cultures or DNA.

The certification approach has been particularly valuable in software development, through the activities of the Open Source Initiative¹⁵ which oversees the branding of such licenses associated with copyright of free and open source software. However, life sciences are extremely sector-specific and technology-specific, and it is impossible to forecast or fully anticipate the emerging patent rights; these facts complicate BIOS certification and licensing. Of course these same challenges also render patent-based BIOS licensing and MTAs even more necessary.

Patent Lens: A Platform for Understanding IP Landscapes

With funding from the Rockefeller Foundation, in 1999 CAMBIA began to develop an integrated, full-text database of patents in the agricultural sciences. Under the initial guidance of Dr Carol Nottenburg, then CAMBIA's Director of Intellectual Property, the CAMBIA IP Resource became a prominent web-based data tool to investigate patents in this field. Over the years, both the ambitions and the capabilities of the CAMBIA Patent Lens team grew,¹⁶ and Patent Lens has now become one of the world's foremost cost-free resources for full-text searching and understanding patents in many jurisdictions and in all classifications. Patent Lens¹⁷ harmonises, parses and presents worldwide patent and technology data in a full-text searchable and highly integrated manner.

However, it is much more than a patent database. Patent Lens is an integrated response to the massive complexity and opacity of the world of patents. It is intended as a public platform to enable many actors to investigate and share analysis of relevant IP issues, and to foster community involvement in overseeing and guiding the patent system.

¹⁵ Open Source Initiative <opensource.org>.

¹⁶ The Patent Lens was featured in an editorial in *Nature Biotechnology* (24 2006 474), called 'Patently Transparent' which was disarmingly positive about our Patent Lens activity providing a critical breath of transparent fresh air to the patent frenzy that is creating a crisis in biotechnology. The Patent Lens team, led for the last two years by Dr Marie Connett, still has its original three software informatics specialists, Greg Quinn, Doug Ashton and Nick Dos Remedios, and has been strengthened by additional talent, including Paul Freeland, Neil Bacon and Josh Cole.

¹⁷ Patent Lens <www.patentlens.net>.

The patent system has grown so rapidly and become so complex and opaque that even the most privileged and skilled clergy of patent law can only parse a tiny area of specialised knowledge, and that tiny area changes daily. This fragmentation has made it almost impossible to thoughtfully and factually assess the consequences of action and inaction: how can the consequences of policy be modelled or validated when patents are treated as fungibles? How can efficient progress in sectors critical to social progress, such as health, environment, and agriculture, be secured when the rights are tangled in a skein of patents?

The goal of the Patent Lens is to use the power of informatics and community to harmonise and make transparent the world of patents, so that thoughtful individuals, institutions and agencies can guide thoughtful and humane reform of the innovation system and to spur efficient and socially relevant innovation. This is an essential platform if we are to make use of the patent system itself to expand and protect a technology commons, and to collectively target breakthrough inventions, work-arounds and ‘work-beyonds’¹⁸ and to make thoughtful and informed partnerships.

BioForge: Field of Dreams?

BioForge was initially launched as a web-based collaboration platform to take CAMBIA’s pump-priming technologies—including Transbacter (described later), a new generation GUS called GUSPlus, and a novel genetic fingerprint technology called DArT—and throw open the gates to enlightened self-interest. We wanted scientists to try Transbacter in diverse bacteria and crops to create an open source and effective toolkit. The first version of the web facility was based on a very credible collaborative software development platform created by Brian Behlendorf¹⁹ and his colleagues at Collabnet. We had hoped—in retrospect, perhaps naively—to see a surge of interest: scientists from around the world, initially from the public sector, would register, log on,

¹⁸ A work-beyond refers to a created technology which both bypasses and transcends the proprietary technology it seeks to replace. Transbacter, described later, is an example of a ‘workaround’, which will become a work-beyond when its efficacy and uptake increases.

¹⁹ Brian Behlendorf is the Chairman of the Apache Software Foundation, and a driving force in the creation of the Apache Web server, one of the most widely used open source software tools in the world, with nearly 70% of the world wide web making use of it.

and offer to collaborate to improve these tools, and to share their thoughts and actions.

The initial response was mildly enthusiastic, but within a few months we realised that the actual engagement and contribution of scientific or personal resources was miniscule. While the BioForge has almost a thousand registered users, very few of them have substantially assisted the listed projects, technically or scientifically. However, many of the registered users are from India, China, and other countries widely viewed as out of the mainstream of cutting-edge biological research. This may reveal a latent need or desire for a better-crafted collaboration culture. We also believe it reflects CAMBIA's reputation as a provider of enabling technology. Thousands of our pCAMBIA DNA vectors toolkits are in use in almost every country, so this 'market' knowledge and confidence could also be skewing the numbers. Still, at this stage BioForge has yet to create a vibrant web-connected community that actually does anything. We use it constantly, as a transparent and inclusive 'lab notebook' for our own work at CAMBIA.

To address the issue of enhancing contributors' reputations (see BioForge textbox), CAMBIA has started a software development project called Karmeleon to create open source, modular, software-mediated reputation metric tools. We hope that people in many collaborative and distributive projects can use these tools, and tune them to their diverse needs, ranging from online review of scientific publications through to research collaboration and product development. Our premise is that individuals should be rated on their contributions by accredited (rated) peers in a transparent manner, but using sophisticated, multivariate metrics to reflect the complex and diverse nature of the value of their contributions. Beyond their professional value, these contributions can and often do have important community and utility implications.

If we make valid, less 'game-able' metrics available, users can develop confidence in the value of one another's contributions, and provide rewards as their community norms dictate: career advancement, peer reputation, funding and so on. But the reputation metrics must be adaptable to the culture where the contributor is working and being evaluated. Our initial drafts of Karmeleon use three metrics: Community value, Utility value, and Professional value. Scores in each category in turn impact the 'gravitas' of a user; we hope this will encourage more sensible ratings to emerge.

BioForge: The Challenge of Aligning Incentives and Rewards

In initially designing BioForge, we had hoped that scientists in public sector institutions would come to see the value of working together to build powerful common toolkits to solve problems. Clearly most public entities endorse and even encourage the notion of pulling together to solve intractable social and economic problems: market failures. Indeed, this is the best justification for the very existence of a public sector. But if the toolkit does not encourage scientists to solve problems for their self-interest, it will be irrelevant. And if such participation carries a cost—in real time and resources—that is yet another disincentive.

Furthermore, while discovery and occasionally invention are activities within the public purview in universities and government agencies, innovation—the delivery of new and tangible improvements to society—is not. Hence it is not part of academic science culture to be aware of the challenges to innovation. Nor does academia do much to reward sharing. The metrics for success are almost always being ‘first’ in a field of endeavour that is widely hailed as being important and timely. The grind of innovation, with its need for long timelines and the building of confidence at many stages of product or process delivery, has little appeal and less relevance to academic advancement. In fact, the market increasingly rewards those who monetise or sequester the necessary components of innovation—a perverse set of incentives if there ever was one. Discoveries are routinely patented; while they are only part of the complex web of capabilities that must be aggregated to create wealth, owners can game them for short-term financial gain at the expense of sectoral progress.

Success with a BioForge project—or any cooperative project with long timelines and complex feedback loops—requires aligning incentives and rewards. The most prominent metric for academic advance is reputation, but the tools for recognising and enhancing reputation are still very primitive, including publication in high-impact peer-reviewed journals and serving on committees and review panels to cement relationships.

BioForge lacks any mechanism to demonstrate its contributors’ influence and success to the community at large, or to those entities and individuals that have power over professional advancement. It takes an

exceptional scientist to work toward improving a technology if she or he has no personal stake in its success.

The long timelines of agricultural and medical research and product development all but forbid direct feedback when an innovation enters the market-place. This is a key justification for vertically integrated companies: to ensure that managerial oversight creates these links. If we wish to see alternative, distributive innovation in sectors with such challenges, we must create intermediate, interconnected and valuable feedback that enhances contributors' reputations, as well as new incentive pulls to participate.

The first generation of BioForge taught us something fairly obvious: that the cultures of software engineering and the life sciences overlap very little. Software developers live online. Their tool—the computer—is their window to the Internet. Their product, software code, can be tested almost instantly and can be evaluated, rejected or accepted almost as quickly. The engineer can build on tested code, and be fairly confident of a secure base. In the life sciences, experiments can take months or years; validation, scaling and quality assurance take even longer. And the process can be so expensive or so specific to circumstances that it may never be replicated by another entity.

We are cautiously optimistic that as we introduce new, recognised and respected 'reputational' tools, if we nurture high profile and energetic champions for particular projects, and if we create new incentive and reward systems, we will be able to move the BioForge from a field of dreams into a productive and focused mechanism for distributive innovation.

An ‘Apollo Project’ for Biological Innovation?

Several months after we published our TransBacter paper in *Nature*, *Nature Biotechnology*—the most prominent scientific journal in the commercial biotechnology sector—published an editorial expressing scepticism that a true open source movement could happen in biotechnology, given the extent of entrenched norms and interests.²⁰ The title of the editorial, ‘Open Sesame’, implied that a vision as clearly utopian and impractical as that of open source for biotechnology would need a magic incantation in order to become reality.²¹ The article did conclude, however, that an open source movement in biotechnology might just take root if, in an ‘Apollo Project’ of some type could be used to forge a common ground to develop new collaboration norms, tools, business models and science around some mutually agreeable and highly desirable goal.²² While we at CAMBIA do not agree with the editors of *Nature Biotechnology* that the only way forward for open source in biotechnology is a grand-scale ‘Apollo project’ of the type they suggested, we do agree that it may be an attractive option.

What would a 21st century Apollo project to spur biological innovation look like? If the BIOS Initiative and the movement need such a platform from which to explore, create and coordinate new modes of problem solving using life sciences, what will that platform be? First, the project would require a socially and economically highly desired goal

²⁰ An outstanding article by Kenneth Cukier appeared about a year later: ‘Navigating the Future(s) of Biotech Intellectual Property’ (2006) 24 *Nature Biotechnology* 249–51. It articulately described the increasing impasse in biotechnology caused by misuse of the IP system, and featured CAMBIA’s BIOS Initiative very prominently and favourably. The metaphor Kenn used in this paper—that of maritime navigation and commerce—is extremely apt and informative. His paper is strongly recommended.

²¹ ‘Open Sesame’ (2005) 23 *Nature Biotechnology* 633. Clearly the authors did not have a young child to remind them that ‘Open Sesame’ was the incantation that would open the cave in which thieves had already sequestered stolen riches, a suitable parable for the misuse of the patent system.

²² The Apollo project was the concerted effort by the United States government to reach the moon before the Soviet Union did. The long-term focus may have been to reach the moon, but the project’s real purpose was to coordinate massive scientific, engineering and technological progress with industrial development, while building and preserving a societal and political confidence associated with success. It wasn’t really about reaching the moon, it was about being able to reach the moon.

for which a technological intervention of great promise can be articulated. The project would need to focus on catalysing new opportunities for problem solving, not just on creating an imposed 'solution'. It would not have a linear impact, nor would it merely improve the cost effectiveness of conventional paradigms.

To engage both the scientific and the business community, such a coordinated effort would offer an intellectually exciting proving ground for new collaborative approaches and new science and must require interdisciplinary skills. The imagination and creative energy of science would be harnessed, but much of science is intensely self-absorbed. An interesting problem will attract much more attention than a mundane one.

The platform activities would afford opportunities for 'spin off' value for other initiatives and activities, and would have impacts beyond its target goals. A broad constituency must see some merit in various components of the project—so that diverse, even divergent interests would build coalitions.

The project would also have a credible promise, or proof of principle.²³ It would not be too risky—or too safe. While it may be somewhat encumbered by intellectual property, it would not yet be completely constrained. If the target has a suite of challenging IP thickets, that would be a platform for new strategies—of decision support, collaboration and invention—to emerge, allowing us to hone these capabilities. It would be in a field with few entrenched interests, or those interests must be diffuse or distracted. If major economic interests push back too early, they could slow or stall the effort.

Finally—and critically—it would also be in an arena where civil society, industry and academia can engage constructively towards a *détente*, and where they can explore and validate new models of social enterprise and business, as well as new economic and innovation strategies.

²³ In the absence of jet aircraft, rocket propulsion and supersonic flight, the idea of space flight would have seemed ludicrous to many.

Beyond the Thicket: Transbacter

By about 2000, my colleagues at CAMBIA and I had seen so much ‘me-too’ science going on around the world and the vast increases in patenting and vertical industry integration. We also saw public support eroding for genetic modification and then for all scientific interventions in agriculture. So we decided it was time to act more aggressively.

We decided to attack the first and most prominent thicket of patent rights—that around *Agrobacterium*—which represented the beginning of the patent rush in agricultural biotechnology. We chose this technology not because we believe that it presents a unique or critical bottleneck to many new entrants into the sector, or because anyone has called for these patents to be revoked or broadly licensed. In fact, these tools have little market pull now. The ‘scorched earth’ policy in the agricultural biotechnology sector has left virtually no inventive entities queuing up to develop products, and no public desire for such products.

Rather, we wanted to show the potential for a new combination: what if we combined patent informatics and transparency with creative, targeted scientific research, and new normative and licensing tools? What if we used it to build a true public commons of technology—or rather ‘rebuild’ a public commons of capability. We sought not a silver bullet, but rather a platform to test and explore our hypothesis that in alternate universes of innovation, tools and foundational discoveries could be constantly improving common goods, and that prosperous industries and business could be built on them.

Assessing the Patent Landscape

In about 2000, we began a comprehensive analysis of the patent situation surrounding *Agrobacterium*-mediated gene-transfer (AMGT), the process I discussed earlier. We intended to publish a simple white paper describing this key thicket of rights. But the task proved much more complex. Ultimately we published the first analysis online; almost 400 pages, and covering the top few hundred patents,²⁴ it has since seen two major updates. Over 1000 users downloaded it. But as we began to realise the extent of the problem, we also realised that it could not be attacked piece by piece. As we analysed the ‘patent landscape’, we noted

²⁴ See <www.patentlens.net>. The first version was mostly a tour de force by Carolina Roa Rodriguez with guidance from Carol Nottenburg.

that all of the patents used a common language and set of definitions that dated to the original filings: that the inter-kingdom gene transfer was achieved as a unique event mediated by a particular bacterial species, *Agrobacterium tumefaciens*.

Definitions are the key to a patent; they are critical in a patent prosecution to establish the metes and bounds of the claimed invention, and to guide courts in the event of a dispute. And the pioneering inventions typically establish precedent that persists. In the case of *Agrobacterium*-mediated gene transfer, it was widely believed and promoted that *Agrobacterium* was a one-off; a unique situation in biology. To this day most scientific papers baldly state that it is the only such situation.

The Strategy

My logic, and that of most biologists trained in evolution, is that if something happens once in life, it probably happens many times—maybe ubiquitously. We think of a ‘one-off’ because we can rarely see other instances. So I began looking for hints in the literature that other bacterial species could transfer genes to plants, either natively or with a bit of convincing. And I found hints aplenty. So we set out—again with support from the Rockefeller Foundation—to find or generate the capacity for benign plant-associated bacteria to conduct gene transfer, and thus to develop a system that would be competent to transfer genes to plants, which was not infringing any *Agrobacterium* patents. If we could do this, the toolkit would clearly fall outside all the patents over AMGT, rendering hundreds, even thousands of patents irrelevant as blocking tools, but useful as ‘background science and technology’.

We further speculated that we would be able to develop a system that was not only free and clear of the onerous *Agrobacterium* thicket, but would ultimately be superior to *Agrobacterium* as a technology. *Agrobacterium* is a plant pathogen, which normally causes disease in susceptible plants. Plants—even non-susceptible ones—seem to know this, and become stressed. We reasoned that by using totally benign symbionts, we’d eliminate the stress on the plant, and open new opportunities for genetic enhancement. If we could make the technology more efficient and wide-acting than *Agrobacterium*, a wholesale migration to the use would occur, even by academics. This

would infiltrate the new open source norms into that most conservative of communities.

The Re&D

The process turned out to be more straightforward than almost anyone expected, and we published our results, which described a new system called ‘Transbacter’, in *Nature*²⁵ on 10 February 2005. After nearly two years of hard work by a skilled laboratory staff, we described in that paper how we had induced three different genera of benign plant bacteria to transfer genes to three different genera of plants.

These plants included the world’s most important crop, rice, over which Japan Tobacco held dominant rights, and broadleaf plants, over which Monsanto held dominant rights.

The capability of *Agrobacterium* to transfer genes to plants is virtually identical at a molecular level to the ubiquitous system by which virtually all bacteria exchange genetic material, and even by which proteins and other molecules are secreted. This similarity allowed us to excise and move this capability on a fairly well-defined DNA construct into the benign symbionts. We were able to test the system with the most sensitive tools in the sector: the open-sourced GUSPlus reporter system.

The paper received exceptional coverage in the press, ranging from the *New York Times* and *Science* to *Nature Biotechnology* and the *Economist*, but not just for its scientific contributions.

The BIOS Licensing Framework

To share this technology, perhaps counter-intuitively, we filed patents on it. At first glance, this is anathema to open sharing. But we were learning the lessons of positive selection and the ugliness of patent gaming and trolling (for an example, see appendix). As we developed the new technology we also developed, in parallel, draft licensing templates for a prototype ‘BIOS’ license, as I described earlier. Two years later, we have over fifty licensees, including large multinational corporations, small companies, and diverse public sector institutions. We have recently stream-lined this technology to be more universal and

²⁵ ‘Gene Transfer to Plants by Diverse Species of Bacteria’ (2005) 433 *Nature* 629–33.

easily disseminated, and have distributed over 300 kits of the new materials. Traction is building as the technology is improving.

But this is not really transformative, merely illustrative and instructive. Real transformation occurs when completely new actors are brought into innovation systems, and when radically new options for problem solving emerge. This is our next ambition.

BIOSENTINELS: A 3D VISION FOR EQUITABLE INNOVATION

The most powerful impact of the scientific method has been to help us understand what had been incomprehensible; it has also helped us visualise and measure the parameters of the natural world. The importance of measurement cannot be over-stated. Without the ability to measure—to see the consequences of an experiment or intervention—we cannot understand it, or improve or build upon it. The future of biological innovation will similarly hinge on turning the unseen into the seen, and to sensibly report on the world around us so we can better respond.

Most critically, we must *democratise* these abilities, both to measure and to respond, in order to *diversify* agro-ecosystems and environments and *decentralise* the problem-solving capability. We will achieve this by fostering scientific method and harnessing local knowledge and commitment in communities that have previously been ignored or treated as passive recipients of help. This is our 3D vision, and the BioSentinel project will be the platform for exploring and realising this vision.

In many vineyards around the world, rosebushes are attractively located at the end of each row. This curious planting regime does not reflect some shared aesthetic among winemakers or grape-growers. Rosebushes are sensitive to certain fungal diseases that affect grapevines more than the grapes themselves. If they plant and observe roses, growers can easily see the early stages of fungal infection on the roses, and can take measures to prevent disease in the grapes. The rose is a natural BioSentinel.

The Role of Measurement in the Next Green Revolution

It is often said—and it is true—that the Green Revolution, which so transformed the agricultural and economic fabric of Asia and much of the rest of the world, passed Africa by. The Green Revolution is not largely about plant breeding, although the short-stature varieties garner great attention. Rather the great advances were in the availability and management of inputs in agriculture. Water, nitrogen, phosphorus, potassium, acidity and countless micro-nutrient and abiotic stresses can each separately and together constitute major production constraints, as well as input costs, to an agricultural system. Combine this complexity with the countless impacts of biotic challenges such as pests and diseases, especially cryptic or latent soil-borne diseases, and creating any kind of profitable and ecologically sustainable farming becomes horrifically complex in the best of circumstances. Little wonder that industrial agriculture's greatest successes—with their concomitant problems—come from homogenising these environments with massive inputs and then breeding and managing these artificial and unstable conditions to get maximum yields.

These options are not available for transforming low-input, low-output agriculture into a prosperous enterprise. When capital, infrastructure and communications are precarious, it becomes even more crucial to accurately and judiciously source and apply suitable nutrition, and to guide management decisions well.

The management of natural resources, whether endogenous or enhanced by inputs, is the most critical and challenging bottleneck in agriculture. It will be the lynchpin of the next Green Revolution. It is also the component most amenable to measurement. But here is the conundrum: to have a sustainable and scalable impact, such management decisions must be made by local problem solvers, and many such people are extraordinarily poor. They cannot afford to measure, and they cannot afford not to.

For the last 15 years CAMBIA has been working on the components necessary to generalise this phenomenon.²⁶ Now, with the advent of

²⁶ This work has benefited particularly from early contributions of Kate Wilson and Steve Hughes, both Members of CAMBIA, now with CSIRO and Exeter University, respectively.

new scientific understanding, new proofs of principle, and the BiOS Framework, this work can now be brought to scale. With initial support from the Lemelson Foundation, we are beginning to create an open source platform to use plants as versatile living BioSentinels to measure and report on the status of their environment.

Imagine a plant—not necessarily a food plant—that has been engineered as an instrument to produce a colour, a smell, or a shape that indicates the level of nitrogen or another essential nutrient in the soil. This plant will be developed in a collaborative, open sourced environment with components that are BiOS licensed and held in public trust. It will be a cost-free instrument that allows any farmer to better judge the condition of her cropping system, and to create wealth by making careful decisions, informed by measurements of the unseen parameters that influence her crop and its environment.

But the BioSentinel project involves much more than engineering one plant to make one colour in a glasshouse. It is no mere academic curiosity. We intend to develop the platform to create a modular toolkit for the public and private sectors alike. We envision mixing and matching components to sense virtually any parameter (nutrient, water, pathogen), transmission of this signal via open standards, and reporting on this parameter with any of several different detection systems (colour, fluorescence, smell, form). We also intend to consider all the quality assurance, regulatory and other parameters necessary for diverse collaborators to create practical and deliverable innovations. The BioSentinels will cost nothing to manufacture, once developed. They will cost nothing to use. But they will add value through the information they make available.

This platform will be built using technologies developed under BiOS license, guided by sophisticated patent informatics to ensure permissive use, and will pioneer new collaborative research methods that enshrine and perpetuate permissive use by all parties. The platform need not create GMO foods, but will create new communities of informed

Summarised in, for example R A Jefferson, 'Beyond Model Systems: New Strategies, Methods, and Mechanisms for Agricultural Research' (1993) *Biotechnology R & D Trends*, in Volume 700 of *The Annals of the New York Academy of Sciences*, 53–73; K J Wilson et al, '...Glucuronidase (GUS) transposons for ecological and genetic studies of rhizobia and other Gram-negative bacteria' (1995) 141 *Microbiology* 1691–705.

decision makers who are empowered to evaluate and improve their own ecologies and economies.

CONCLUSION

At the start of the twenty-first century, science is at a critical juncture. Four centuries of inquiry, discovery, and invention have created a base of knowledge that has the potential to provide people everywhere, in all circumstances, with nourishment, improved health, and longer life. But the institutional mechanisms that ostensibly exist to encourage the application of science to practical problems are today hindering that very process. The norms that have evolved around gate-keeping have created new clergy, new impediments and new inefficiencies. Without asystemic change, science's promise will not be available for those who most need it, and the promise of a truly diverse, robust and fair innovation culture may elude us.

Patents are at the heart of the system of institutions that convert basic scientific knowledge into practical applications. The modern patent system was intended to advance the public good by balancing the disclosure of ideas and the transparent definition of limited property rights. Today, it has degenerated into an instrument that is often misused to obstruct the public good through enclosure of ideas and obscure assertion of property rights that have no concomitant social benefit. To the shared dismay of both scientists and thoughtful citizens, patent systems and the myriad gaming practices they have spawned today are impeding innovation as a social enterprise, and continuing to deprive most of the world's population of such fundamentals as adequate nutrition, access to health care services, and clean water. This does not have to be. It is up to us to reclaim the beauty of science as a democratised tool for social advancement and wealth creation. It is up to us to write the terms of the compact. It is up to us to move beyond rhetoric and into constructive engagement in reforming our innovation systems for economic robustness and social justice.

APPENDIX. CO-OPTING THE COMMONS: A NEGATIVE EXPERIENCE OF POSITIVE SELECTION

For nearly seven years, with expenditures of over \$100 000, CAMBIA has battled Syngenta, the large Swiss agribusiness, in European Patent

Office opposition proceedings and appeals over the validity and scope of Syngenta's patents on 'Positive Selection'. These broad patents (e.g. EP 601092, but with counterparts in the USA) were granted with sweeping claims that conferred on Syngenta an absolute monopoly on 'positive selection' in plants.

Positive selection is the provision of a benign compound—such as a sugar—that an organism cannot use without the action of a new gene; thus it 'selects' for those organisms that have acquired that gene. Positive Selection is one of the most basic tools in genetics, used since the beginning of microbial genetics; all the bacterial genetics in the 1950's and 60's was based on one bacterial strain gaining the ability to grow on new sources of carbon and energy. When I started working with plants, it was thus immediately obvious to me (and presumably to anyone not employed at the patent office) that we could easily adapt this concept to plant genetics, to determine when a new gene had been added to a crop plant, and that a good first use would be my GUS gene.

So I began adapting GUS for this purpose, around the time I started sending out GUS kits and information, and giving hundreds of lectures on its use. While this mode of distribution was to dramatically change the field, it also allowed some aspects of the system to be co-opted. Our ideas and hard work were basically turned from 'non-rival' goods that were available for all as we intended, into a private monopoly that could, and did, suppress innovation by competitors.

Scientists at a Danish sugar company, DANISCO, filed a patent well after I had given them the GUS gene, and after I had given public lectures on the use of GUS for such purposes. In this patent, they were granted broad claims to all uses of positive selection, with any compound and any gene in any plant. This breathtaking scope of claims was based solely on experiments described in the applications that used the GUS gene to activate a biological compound that would allow plant cultures that had GUS to stay green and be 'selected'. This was fundamentally what I had already reported at international meetings, with data showing that it worked. Like many scientists, when I reported it at international congresses, I intended to see it shared with everyone. DANISCO's intention clearly was not.

The potential value of this patent estate caught the eye of Heinz Imhof, then chairman of Novartis, who intervened personally to buy the patent

applications from DANISCO outright. These patents then served as powerful ammunition in the patent war chest of Novartis, which went on to merge with other companies in the vertical integration frenzy of agricultural biotech, to become Syngenta. The evolving strategy of Mutually Assured Destruction by Patent Estate between the large multinationals required just such weapons.

The breadth of the claims as granted in Europe—together with their counter-parts in the USA—ensures that any entity using the approach of conferring a growth advantage on a cell or plant to obtain transgenic plants would be infringing. This left only the use of antibiotic resistance and herbicide resistance as the means of selecting transformed plants. The adverse public response to such antibiotic gene use is well documented.

Thus the environmentally attractive and benign technology of cleaving a sugar and growing preferentially, with no antibiotics, was denied to the world's agriculture community by one group of patents, whose entire rationale was derived from work that I had intended to make public. But with the patent, it was 'enclosed'.

I had several meetings with Imhof and others at Syngenta; I attempted to make the case that using GUS to garner such a powerful and oppressive patent position was unjust and inappropriate and would ultimately be a pyrrhic victory for the sector. The discussions went nowhere.

So we made use of one of the few remedies afforded in the patent system to small players: the opposition process. Once patents are granted in Europe, they can immediately be challenged if one submits to the European Patent Office (EPO) prior art that had not been considered. Our contention in the EPO was that much public work, as well as my own work, including my public disclosure of the basic idea, pre-dated the filings and would thus invalidate the novelty requirement for the patent. We also argued that the patent was obvious in light of the pervasive use of positive selection in every other biological system for many years. We also asserted that the patent did not sufficiently enable one to practice the invention, and in particular, did not merit the breadth of claims granted.

The opposition process is widely touted as much more affordable than litigation. No doubt this is true. Instead of paying several million

dollars to lawyers so we could be screwed by a multinational corporation in front of a judge, we only had to pay a hundred thousand or so for the same privilege, but in front of a panel of patent professionals. Of course reconsiderations of patent validity are conducted by the very same entity—the administrative machine of the patent office – that made the initial patent grant. So even in the face of what we felt to be compelling prior art, and convincing case law, the deck was stacked in favour of the status quo.

Watching the process, and the craft and gaming skills involved, was an eye-opener for me. Until one has actually endured the multi-year posturing, arguing, heartache and expense, there can be no clear way to convey the dysfunction of the system, or its debilitating effect on inventors. We achieved only modest inroads in restricting the breadth of their claims. But we did consume years of time and huge amounts of money, in a failed bid to restore for public use a key application of a technology that I had developed and had inadvertently let a multinational pull into its private fiefdom. The opposition process is not available in the United States, so the opportunity to lose extravagant sums of money there was denied to us.

What did Syngenta do with this technology? With the example they claimed using GUS, nothing. They never made a single product using that tool, nor did they develop it further. But they used the broad claims, granted by both the European and U.S. patent offices, to ensure that no other player—large or small—attempted positive selection without becoming beholden to them. Later, from DANISCO, they acquired other examples of positive selection protocols which worked pretty well and were protected under the umbrella of the broad claims, they made them ‘available’ under a research license to unsuspecting scientists in the public sector. This ‘research license’ strategy is one of the most pernicious co-opting approaches used by large private-sector companies. Once a tool is used under such a license, the only way to then release a product is through after-the-fact negotiations for a ‘commercial license’. Several friends have gone through this process and reported a bare-knuckled strategy that gives the licensee almost no share in the benefit of the product they developed. Few takers, of course.

What are the lessons. Don’t share? This is not a lesson I cleave to, nor a recipe for social progress. Could it have happened otherwise? Absolutely. This example was a case study of how ‘open source’ licenses

could be crafted and protect the public commons, yet allow the private sector to build prosperous businesses using that commons of technology. Perhaps I should have only sent the GUS gene and disclosed the information to those who agreed to terms by which they would share improvements that specifically used GUS; then the entire broad positive selection concept would likely have stayed available to all entities—public and private, large and small—that wished to explore its use. As would the many modifications on which others had filed patents. Just imagine: what would happen if the public sector technology transfer professionals had access to such a leverage tool to further the power of the commons toolkit and advance their mission?

PART EIGHT

INTERNATIONAL DIMENSIONS

e-RESEARCH AND JURISDICTION

Gaye Middleton¹

As part of their daily activities, those involved in e-research will often transfer information, including background materials, research results and software, across state and national borders. The act of transferring information across state and national borders raises a number of jurisdictional issues. This chapter will discuss key issues regarding intellectual property, privacy and dispute resolution as they arise from e-researchers transferring information across state and national borders, and how these issues may contractually be resolved.

INTELLECTUAL PROPERTY ISSUES

There are two key jurisdictional issues relating to intellectual property rights which are raised by the transfer of information across state and national borders. These are:

- the differing intellectual property protection available in different jurisdictions; and
- the possibility that the transfer of information or materials from one jurisdiction to another may result in intellectual property infringement in the recipient jurisdiction.

These jurisdictional issues arise because intellectual property rights are territorial rights, as they are created by national laws that, subject to certain limited exceptions, only apply within the boundaries of the nation which passed those laws.

¹ Inhouse Counsel, Uecomm Operations Pty Ltd.

Differing Intellectual Property Protection

E-researchers engaged in cross-border research projects may be affected by jurisdictional differences in the protection afforded to intellectual property rights.

There are numerous international treaties, such as the *Berne Convention for the Protection of Literary and Artistic Works* and the *Agreement on Trade Related Aspects of Intellectual Property (TRIPS)*, which require the signatories to enact intellectual property laws which comply with the standards laid down by those treaties.² However, even as between member states to these treaties, intellectual property laws are not uniform. Factors which contribute to this lack of uniformity between intellectual property laws in countries which are member states of the same international treaties include:

- differences in the wording of intellectual property legislation between those member states; and
- differences in the way in which those treaties and the implementing legislation is interpreted by the courts of those member states.

For example, in Australia, an author is entitled to copyright protection for a work that the author has created if it originated from the author in that it was not copied, and the author expended skill, labour or expense in creating that work.³ No degree of inventiveness or originality is needed to obtain copyright protection for a work in Australia. This means that a compilation of unoriginal facts or figures may be protected by copyright in Australia if the author can show that labour, skill or expense were used in making that compilation.⁴ By contrast, in the United States, both labour and creativity are required to obtain copyright

² See generally A Fitzgerald and B Fitzgerald, *Intellectual Property in Principle* (2004).

³ *University of London Press Ltd v University Tutorial Press Ltd* [1916] 2 Ch 601, 601–9 (Joyce J); *Victoria Park Racing and Recreation Grounds Co Ltd v Taylor* (1937) 58 CLR 479, 498 (Latham CJ), 511 (Dixon J); *Desktop Marketing Systems Pty Ltd v Telstra Corporation Limited* [2002] FCAFC 112, [160]; *Nine Network Australia Pty Ltd v Ice TV Pty Ltd* [2008] FCAFC 71.

⁴ *Ladbroke (Football) Ltd v William Hill (Football) Ltd* [1964] 1 WLR 273 (HL), 285 (Lord Hodson), 289 (Lord Devlin); *Computer Edge Pty Ltd v Apple Computer Inc* (1986) 161 CLR 171, 182–3 (Gibbs CJ); *Desktop Marketing Systems Pty Ltd v Telstra Corporation Limited* [2002] FCAFC 112, [160].

protection for a work.⁵ This means that works such as a database of facts and figures compiled by researchers may be protected by copyright in Australia, but when that database is transferred by researchers to another jurisdiction, such as the United States, it may not be protected by copyright.

Further, the duration of copyright protection varies between nations. As a result of the implementation of the *Australia – United States Fair Trade Agreement*,⁶ Australia and the United States both provide general copyright protection in works for the life of the author plus 70 years after the author's death.⁷ By contrast, in Japan, the term of general copyright protection in works is currently life plus 50 years.⁸ This means that works that are no longer protected by copyright and hence are able to be used freely by researchers in one jurisdiction may remain subject to copyright protection in another jurisdiction. For example, if researchers in a multi-jurisdictional research collaboration need to use the same work, the researchers in one jurisdiction may be able to do so without restriction, while the researchers in another jurisdiction may have to obtain a licence of that work for the purposes of their research.

Transfer of Information and Intellectual Property Infringement

Another jurisdictional issue relating to intellectual property rights which faces e-researchers is that, due to the differences in intellectual property laws between countries, the transfer of information or materials from one country to another may, in certain circumstances, result in intellectual property infringement in the recipient jurisdiction. This infringement may occur even though the use of the transferred information or materials in the originating jurisdiction does not infringe the intellectual property rights of others. The reason for this is that, although intellectual property rights are territorial in nature, there are certain circumstances in which the intellectual property laws of a country

⁵ *Feist Publications, Inc v Rural Telephone Service Co Inc* 499 US 340 (1991).

⁶ For a copy of the Free Trade Agreement and associated documents, see: <http://www.dfat.gov.au/trade/negotiations/us_fta/final-text/index.html> at 16 January 2007.

⁷ *Australia - United States Free Trade Agreement 2004* art 17.4.4; *Copyright Act 1968* (Cth).s 33.

⁸ *Japanese Copyright Act* art 51.

can apply extraterritorially. This means that acts done outside of that country may infringe the intellectual property laws of that country.

One example of the extraterritorial application of Australian copyright law is liability for authorisation infringement. A person who uploads a work protected by Australian copyright onto the Internet outside Australia without the author's authority could be liable for authorising infringement of copyright in that work within Australia if an infringing copy of that work is made in Australia from the unauthorised copy that was uploaded onto the Internet outside of Australia. This is because, if a person authorises an infringing act by another person that takes place within Australia, it does not matter for the purposes of section 36 of the *Copyright Act 1968* (Cth), which deals with authorisation infringement, that the first person authorised the infringement outside of Australia. The fact that the infringing act which that person authorised takes place in Australia is sufficient to establish liability for authorisation infringement.⁹

Some United States courts have applied United States copyright laws outside of the United States by applying a principle known as the 'root copy' doctrine. Under the root copy doctrine, a person may be liable for infringement of US copyright in respect of a work where that person makes an unauthorised copy of a work protected by copyright in the United States, not only for making the unauthorised copy itself, but also for any distribution of that infringing copy outside of the United States.¹⁰ An example of the potential application of the root copy doctrine in the context of e-research is where a researcher in the United States makes an unauthorised copy of research results that are protected by copyright, and provides that copy to his or her colleagues in Australia, where it is further distributed. In these circumstances, the researcher based in the United States may be liable for copyright infringement both for making the initial copy and for its distribution outside of the United States.

⁹ S Ricketson and C Creswell, *The Law of Intellectual Property: Copyright, Designs and Confidential Information* (2002) [16A.145].

¹⁰ *Update Art v Modin Publishing Ltd* 843 F 2d 67, 82 (2nd Cir 1988). See further discussion of this doctrine in G Austin, *Private International Law and Intellectual Property Rights – A Common Law Overview* (WIPO, 2001) [31]
<http://www.wipo.int/edocs/mdocs/mdocs/en/wipo_pil_01/wipo_pil_01_5.pdf> at 12 January 2007.

In respect of patents, a person in Australia who electronically uses an Australian patented invention hosted by a foreign person on a server situated outside Australia, or who imports data generated by that invention into Australia via the internet, may infringe the Australian patent for that invention. However, the person who operates the website from which the invention is accessed may also be liable for contributory patent infringement under section 117 of the Australian *Patents Act* for supplying a product, in this case, being data or information, which infringes an Australian patent.¹¹

E-researchers may infringe United States patent laws by acts done outside of the United States under section 271(f) of the United States *Patent Act* by supplying in or from the United States, without the patent holder's authority, all or a substantial number of the unassembled components of a patented invention so as to actively induce the combination of those components outside the United States in a way which would infringe the patent for that invention if those components were combined within the United States.

A recent series of cases in the United States considers whether software source code is a component of a patented invention such that, if a person supplies that source code to others outside of the United States, and thereby enables the conduct of foreign activities that would have infringed copyright in the relevant invention in the United States, then the supplier is liable for inducement infringement under section 271(f) of the United States *Patent Act*.

In the first of this series of cases, *Eolas Technologies Inc v Microsoft Corp*,¹² the Court of Appeals for the Federal Circuit held that Microsoft was liable for inducement infringement of an Eolas patent when it shipped master disks containing source code for its Internet Explorer® computer program to hardware manufacturers outside of the United States, who then used those master disks to load Internet Explorer® directly onto their hardware. The Court held that source code can be a

¹¹ J Swinson and G Middleton, 'The effectiveness of patent protection for e-commerce technologies' (Proceedings of the Technology Transfer and Innovation Conference 2001, Brisbane, 2001). This area of patent law is still uncertain, as is the application of *Patents Act 1990* (Cth) s 117. For a summary of decisions regarding the operation of the *Patents Act* s 117, see commentary in J Lahore et al, *Patents, Trade Marks and Related Rights* (Butterworths, subscription service) [18,285].

¹² 399 F 3d 1325 (Fed Cir 2005).

component of a computer program invention for the purposes of section 271(f) of the US *Patent Act*, even though:

- neither source code nor a computer program is a physical object; and
- the content of the master disks was copied onto hardware and the disks themselves did not form part of the final product.

By contrast, in *Microsoft Corporation v AT&T*,¹³ the Supreme Court of the United States found that Microsoft was not liable for inducement infringement under section 271(f) of the US *Patent Act* where Microsoft supplied master disks for its Windows® computer program to manufacturers outside of the United States, and those foreign manufacturers loaded the content from copies of the master disks onto their foreign-made hardware, rather than from the master disks themselves. As Microsoft had not supplied the foreign-made copies of the master disks that the foreign manufacturers combined with their hardware to form AT&T's patented invention, the court found that Microsoft did not supply components of AT&T's patented invention from the United States which were combined to make the patented invention outside of the United States.

These cases are relevant to researchers who electronically distribute software source code from the United States to colleagues in another jurisdiction. They demonstrate that supplying source code which constitutes a 'component' of a patented invention to a person outside of the United States may result in a researcher inadvertently infringing a United States patent.

PRIVACY

Privacy has been defined as the claim of individuals, groups or institutions to determine when, how and to what extent information about them may be communicated to others.¹⁴ There are various categories of privacy, including bodily privacy, privacy of

¹³ US Supreme Court, No. 05–1056, 30 April 2007, overturning the decision of the Court of Appeals for the Federal Circuit in *AT&T v Microsoft Corporation* 414 F 3d 1366 (Fed Cir 2005).

¹⁴ A F Westin, *Privacy and Freedom* (1967) 7.

communications and territorial privacy.¹⁵ However, in this chapter, the term ‘privacy’ refers solely to information privacy, being the rules which govern the collection and handling of personal information. ‘Personal information’ is defined in section 6 of the Australian *Privacy Act 1988* (Cth) as:

Information or an opinion ... whether true or not, and whether recorded in a material form or not, about an individual whose identity is apparent, or can reasonably be ascertained, from the information or opinion.

Within Australia, e-researchers may be bound, not only by the Commonwealth *Privacy Act*, but also by state and territory laws dealing with privacy. While the state and territory privacy legislation is similar to the Commonwealth privacy legislation, there are certain differences. This may raise jurisdictional issues when information is transferred within Australia between researchers situated in different States and Territories. For example, in respect of personal information concerning deceased persons, the Commonwealth *Privacy Act* only applies to persons who are alive, as does the Victorian *Information Privacy Act 2000*. However, under the Northern Territory’s *Information Act 2002*, information about a person continues to be protected for 5 years after their death; in Tasmania, under the *Personal Information Protection Act 2004*, personal information is protected until 25 years after a person’s death; and under the New South Wales *Privacy and Personal Information Protection Act 1998*, personal information is protected for 30 years after a person’s death.¹⁶ Accordingly, even within Australia, researchers must take care to comply with the privacy legislation in each state or territory where that information is transferred.

The Commonwealth *Privacy Act* imposes restrictions on the transborder flow of data under the National Privacy Principles. The National

¹⁵ B Fitzgerald, A Fitzgerald, G Middleton, YF Lim and T Beale, *Internet and E-commerce Law: Technology, Law and Policy* (2007) [10.10], citing D Banisar, *Privacy and Human Rights 2000: An International Survey of Privacy Law and Developments* (Electronic Privacy Centre and Privacy International) <<http://privacyinternational.org/survey/phr2000/overview.html>> at 14 May 2007.

¹⁶ D Lindsay, A Monotti, M Paterson and A Chin, *Legal Issues in eResearch: Report for the Content and Rights Work Package (CR6) – Dataset Acquisition, Accessibility and Annotation e-Research Technologies Project* (2006) 101, Table 5.2 <<http://www.dart.edu.au/workpackages/cr/cr6.html>> at 29 July 2007.

Privacy Principles ('NPPs') primarily apply to researchers in private sector organisations with an annual turnover of more than \$3 million per annum or researchers working for health service providers.¹⁷ This means that researchers in the public sector (that is, government) or in the majority of universities, which are established under State law rather than Commonwealth law, will not be bound by the NPPs.¹⁸ Under NPP 9 in Schedule 3 of that Act, an organisation may not transfer personal information to someone in a foreign country that does not have a comparable information privacy scheme to Australia, unless the individual whose personal information is being disclosed consents to that transfer, or where, among other things:

- the organisation reasonably believes that the recipient of the information is subject to a law or other instrument which requires the recipient to handle personal information in a similar way to the National Privacy Principles; or
- the transfer benefits the individual, and is necessary for the performance of a contract between the individual and the organisation; or
- the transfer benefits the individual, and while the individual's consent cannot be obtained, the transferring organisation can show grounds for the belief that the individual would give their consent if it were possible to obtain it; or
- the organisation has taken reasonable steps to ensure that personal information that it transfers will not be held, used or disclosed by the recipient in a manner which is inconsistent with the National Privacy Principles.

However, NPP9 does not prevent the transfer of personal information outside of Australia by an organisation to:

¹⁷ *Privacy Act 1988* (Cth) s 16A. See also A Hayne, 'Privacy Regulation and e-Research' (Paper presented at the Legal Framework for e-Research Conference, Gold Coast Queensland, 11–12 July 2007) 4 at <http://www.privacy.gov.au/news/speeches/spp07_07.pdf> at 24 June 2008.

¹⁸ The privacy obligations of Commonwealth government departments and agencies are set out in the Information Privacy Principles ('IPPs') in section 14 of the Privacy Act 1988 (Cth). The IPPs do not contain specific provisions restricting transborder data flows. Researchers in universities and institutions established under State law are subject to the privacy legislation in the relevant State. Most States which have legislative privacy principles have based them on the Commonwealth IPPs, hence they also do not contain specific provisions restricting transborder data flows.

- another part of the same organisation; or
- the individual to whom that information relates.

NPP 9 is based on the restrictions on cross-border transfers of personal information laid down by European Commission Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data ('Directive'). Article 25 of the Directive prohibits the transfer of personal data to countries which do not provide adequate levels of data privacy. The European Union ('EU') does not yet recognise Australian privacy laws as adequate for this purpose. One reason for this is that NPP 9 permits the transfer of personal information across national borders where the transferring organisation 'reasonably believes' that the recipient is bound by a law or other instrument which is substantially similar to the National Privacy Principles; by contrast the Directive requires that the recipient *must* be in a country with an adequate level of protection. Another reason why the EU does not recognise Australian privacy laws as adequate for the purposes of the Directive is that it permits an organisation to transfer personal information across national borders where it has simply taken 'reasonable steps' to ensure that the recipient will not deal with that information inconsistently with the National Privacy Principles. This exception is regarded as weak and imprecise.¹⁹

Where personal information is transferred across national boundaries, the differences in privacy laws become significant, making compliance with those laws difficult. To add to the complexity, the United States does not have a single piece of legislation which comprehensively regulates information privacy in the private sector. Instead, information privacy is regulated by a patchwork of sector-specific legislation, the tort of invasion of privacy and trade practices legislation.²⁰

¹⁹ Office of the Privacy Commissioner, *Issues Paper – Review of the private sector provisions of the Commonwealth Privacy Act 1988* (Australian Government, Canberra, 2004) 22–3 <<http://www.privacy.gov.au/act/review/ispap2004.pdf>> at 29 July 2007; Australian Law Reform Commission, *Issues Paper 31 - Review of Privacy* (Australian Law Reform Commission, 2006) [13.59]–[13.72] <<http://www.austlii.edu.au/au/other/alrc/publications/issues/31/>> at 29 July 2007.

²⁰ For example, see J Reidenberg, 'Privacy Wrongs in Search of Remedies' (2003) 54 *Hastings L.J.* 877, 880–1.

Even between EU countries, there is a wide degree of variance in data privacy laws. This arises primarily because of the different ways in which data privacy is protected in the member states. Some countries, such as Portugal and Spain, have enacted specific data protection provisions in their constitutions. Other countries derive data protection from their existing constitutional principles without specifically referring to data protection. For example, Germany derives the right to data privacy from the general right to respect for one's personality. In the United Kingdom, which does not have a written constitution, the right to data privacy is derived from certain rights in the European Convention on Human Rights, which is incorporated into its national laws.²¹

Because of the wide variations in privacy laws between jurisdictions, it can be a compliance nightmare for researchers who transfer personal information across national borders. The Directive regulating privacy in the EU currently sets the highest standard for information privacy; accordingly, if researchers take steps to comply with the Directive in respect of the transfer of personal information across national borders, then they will be likely to have satisfied the privacy laws of most jurisdictions.

DISPUTE RESOLUTION ISSUES

If a dispute arises between researchers in different states, territories or countries, there are two key jurisdictional issues which arise. These issues are:

- first, which courts have the authority to require the parties to the dispute to appear before it for the purposes of deciding the dispute; and
- second, which country's laws should be applied to determine the dispute.

²¹ D Korff, *EC Study on Implementation of Data Protection Directive – Comparative Study of National Laws* (Human Rights Centre, 2002) 8–9
<http://ec.europa.eu/justice_home/fsj/privacy/docs/lawreport/consultation/univessex-comparativestudy_en.pdf>.

Personal Jurisdiction

If a court has authority to require a person to appear before it in respect of a dispute, that court is said to have personal jurisdiction over that person.

At common law, for a court to have personal jurisdiction over a person, that person must be served with court proceedings while that person is physically within that court's geographical jurisdiction.²² Under Australian law, there are limited circumstances in which a person may be served with court proceedings when they are outside the court's geographical jurisdiction. These circumstances, which are governed by the rules of that court, include:

- where the dispute concerns a contract which was made or broken within the jurisdiction, or which is governed by the laws of that jurisdiction;
- where the dispute concerns a tortious act committed within the jurisdiction; or
- where the dispute concerns a breach of legislation which took place within the jurisdiction.²³

This means that a researcher involved in a multi-jurisdictional research project who communicates with the other participants in the research project entirely by electronic means may be subject to court proceedings in another state or country, even if that researcher has never physically entered that state or country.

Choice of Law

Where a court has personal jurisdiction over the parties to a dispute, but where more than one forum has a connection with the dispute, the court must decide which forum's laws it will apply to decide the dispute. To decide which forum's laws it will apply in these circumstances, a court

²² *Laurie v Carroll* (1958) 98 CLR 310.

²³ For a discussion of the circumstances in which a person may be served with proceedings when they are physically outside the jurisdiction of a court, see P Nygh and M Davies, *Conflict of Laws in Australia* (7th ed, 2002) 51–75; Butterworths, *Halsbury's Laws of Australia*, 85 Conflict of Laws [85–335]ff; Law Book Company, *The Laws of Australia*, 5 Civil Procedure [4]ff; B Fitzgerald, A Fitzgerald, G Middleton, YF Lim and T Beale *Internet and e-commerce law: technology, law and policy* (2007), n13 at [2.40]–[2.60].

will use what are known as ‘choice of law’ rules. Accordingly, those involved in e-research may not only be subject to court proceedings in a place other than their home state or country, but they may also be subject to the laws of another state or country in respect of those proceedings.

Different choice of law rules apply to different categories of dispute. In respect of contractual disputes, if there is no governing law clause or other choice of law which can be implied from the contract itself, then Australian courts will apply the law with which the contract has the closest and real connection at the time it is formed.²⁴ Factors relevant to determining which forum’s laws should apply in respect of a contract dispute include the form and legal language of the contract, the place where the contract is made, the place where the contract is to be performed and the place of residence of the parties to the contract. By contrast, Australian courts will apply the law of the place of the wrong, known as the ‘lex locus delicti’, when deciding disputes concerning tortious acts, such as negligence or defamation.²⁵

CONTRACT AS A MEANS OF RESOLVING JURISDICTIONAL ISSUES

The jurisdictional issues confronting e-researchers that are discussed in this chapter can be addressed by a research collaboration agreement between the researchers who are working together on an e-research project.²⁶

Jurisdictional issues which can be dealt with by e-researchers in a research collaboration agreement include the following:

- The agreement may state the laws of the forum which will apply when interpreting and administering that agreement.²⁷

²⁴ *Oceanic Sun Line Special Shipping Co Inc v Fay* (1988) 165 CLR 197, 217.

²⁵ *John Pfeiffer Pty Ltd v Rogerson* (2000) 203 CLR 503; *Regie National des Usines Renault SA v Zhang* (2002) 210 CLR 491.

²⁶ B Fitzgerald and J Abbot, *Legal Framework for e-Research* (2006) 3
<http://eprints.qut.edu.au/archive/00005311/01/5311_1.pdf>.

²⁷ For limits on this approach see B Fitzgerald, A Fitzgerald, G Middleton, YF Lim and T Beale, *Internet and e-commerce law: technology, law and policy* (2007) [2.200], [2.230]. See also *Bragg v Linden Research, Inc and Rosedale* 487 F. Supp. 2d 593 (E.D. Pa. 2007), where Robreno J of the US District Court for the Eastern District of Pennsylvania rejected the defendant’s motion to

- The agreement may state the courts of the forum which will have power to decide disputes arising in respect of the research project.²⁸
- The agreement may state how intellectual property rights developed during the course of the research project will be owned and how they may be dealt with by the participants in the project.²⁹
- The agreement may state the privacy obligations of each party to the research collaboration.³⁰

dismiss the plaintiff's action brought in that court and the defendant's motion to compel arbitration. The defendant's motions cited, among other things, its Terms of Service for using Second Life, its online virtual world. These Terms of Service included a choice of law and jurisdiction clause which stated that the laws of California governed the contract between the parties, and that all disputes would be settled by binding arbitration in San Francisco, California, according to the rules of the International Chamber of Commerce. In registering to use Second Life, the plaintiff agreed to these Terms of Service. However, Robreno J held (at page 611) of the judgment that the choice of law and jurisdiction clause was not enforceable against the plaintiff because:

Taken together, the lack of mutuality, the costs of arbitration, the forum selection clause, and the confidentiality provision that Linden unilaterally imposes through the [Terms of Service] demonstrate that the arbitration clause is not designed to provide Second Life participants an effective means of resolving disputes with Linden. Rather, it is a one-sided means which tilts unfairly, in almost all situations, in Linden's favor.

²⁸ For limits on this approach see B Fitzgerald, A Fitzgerald, G Middleton, YF Lim and T Beale, *Internet and e-commerce law: technology, law and policy* (2007) [2.70], [2.85]-[2.90]. See also the discussion regarding *Brugg v Linden Research, Inc and Rosedale* 487 F. Supp. 2d 593 (E.D. Pa. 2007) at footnote 27 above.

²⁹ However, where researchers develop intellectual property based on a third party's pre-existing intellectual property, the terms of the licence to use that intellectual property may stipulate who owns any intellectual property developed from the licensed intellectual property. For example, it is common for licensors of intellectual property to stipulate as a licence condition that any modifications or improvements to the licensed intellectual property developed by licensees will be owned by the licensor.

³⁰ The terms of a research collaboration agreement cannot override legislative privacy obligations. However, contractual privacy provisions can serve to make participants in an e-research project aware of their statutory privacy obligations, and may also serve as a means to obtain consent from the participants in a research project to the use of their own personal information for the purposes of that project.

A research collaboration agreement which deals with these issues is an important tool for researchers to minimise their potential legal risk. However, a research collaboration agreement is not the only source of their rights and obligations in respect of these issues, and it is vital for e-researchers to take account of other sources of these rights and obligations which may impact on their research project. In addition to statutory rights and obligations, there may, for example, be licences applicable to certain materials used in the research project. These materials may be licensed by traditional intellectual property licences, or be subject to an open access licence. Ideally, e-researchers should identify these other sources of rights and obligations prior to drafting a research collaboration agreement in respect of their research project so that the terms of that agreement are consistent with and subject to those other sources.

CONCLUSION

This chapter discusses key jurisdictional issues which arise from the transfer of research information across state and national boundaries in the conduct of e-research. To minimise their legal exposure resulting from these jurisdictional issues, it is important for e-researchers to enter into a research collaboration agreement which adequately addresses each of these issues. Such an agreement cannot provide e-researchers with an absolute solution to complex jurisdictional issues, given the multiple sources of rights and obligations which may impact on a research project. However, it enables the parties to a multi-jurisdictional research project to better understand the legal obligations that apply to them as a result of transferring information across national boundaries, and to provide themselves with some degree of certainty regarding the applicable laws and forum if a dispute arises between them at a later stage.

PART NINE

SURVEY OF STAKEHOLDER
ATTITUDES

AUSTRALIAN SURVEY ON LEGAL ISSUES FACING e-RESEARCH

Maree Heffernan and Scott Kiel-Chisholm¹

INTRODUCTION

The Legal Framework for e-Research Project lead by Professor Brian Fitzgerald and hosted by the Queensland University of Technology (QUT) is funded by the Australian Commonwealth Department of Education, Employment and Workplace Relations (DEEWR), formerly Education, Science and Training (DEST), under the Systemic Infrastructure Initiative (SII), Research Information Infrastructure Framework for Australian Higher Education, as part of the Commonwealth Government's *Backing Australia's Ability – An Innovation Action Plan for the Future (BAA)*.

The Project involves mapping out a sophisticated legal framework for e-Research and collaborative innovation. As we transition into the National Collaborative Research Infrastructure Strategy (NCRIS)² era it

¹ Project Officer and Project Manager, Legal Framework for e-Research Project (respectively)

This chapter is based on Maree Heffernan and Nikki David, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results* (2007)
<<http://eprints.qut.edu.au/archive/00008865/01/8865.pdf>>.

The authors wish to acknowledge the assistance of Professor Brian Fitzgerald, Dr Amanda McBratney, Dr Anne Fitzgerald and Dr John Abbot and thank them for their efforts in developing and promoting the survey document.

The authors would also like to thank the following people for their valuable contribution to the development and distribution of the survey: Nikki David, Shane Dalglish, Amy Barker, Tanya Butkovsky, DVC Professor Tom Cochrane, Dr Terry Cutler, Professor Mary O'Kane, Margot Bell, Professor Ian W. Turner, Ruth Bridgstock, Professor Paul Roe, Michael McArdle, Kerrin Anderson, Malcolm McBratney, Dr Evonne Miller, Steve Matheson, Dr Graeme Kernich, Dale Gilbert, Ray Duplock, Michael Dean, Mike Finney, Associate Professor Gillian Hallam, Clare McLaughlin, Professor Mark Perry, Terry Bell, Ruth Bridgstock, Associate Professor Chris Collet, Dr Joe Young, Karen Barnett, Dr Vladimir Likic, Professor Bernard Pailthorpe, Professor Stuart Cunningham, Professor Zee Upton, Samantha Cobb, Gaye Middleton and Professor Amanda Spink.

is vitally important that social and legal aspects of the e-Research framework are developed in step with the rapid advances in technology. Only little work has been done in this area worldwide.

This project is linking with key international actors to provide an internationally significant project. While the Open Access to Knowledge (OAK) Law Project³ aims to examine the role of open access to all in an Internet world, this project also focuses on open innovation within secure knowledge communities – both are vital aspects of the e-Research framework. The critical issue is working out legal models for e-Research that reflect the capacity of the technologies involved and can be implemented quickly, effectively and (in many instances) in an automated way.

The Australian Federal government has implemented the National Collaborative Research Infrastructure Strategy (NCRIS) to provide greater focus and strategic direction for Australia's research infrastructure. The NCRIS *Strategic Roadmap*⁴ identifies priorities for investment in research infrastructure. In addition to 15 specific areas of science and technology, 'Platforms for Collaboration' are also designated as a priority capability area.⁵ In addition to hardware and software elements, this priority area includes copyright and other legal considerations.

The conduct of research and the dissemination of its outcomes are greatly enabled by recent and continuing development in communications networks, information and computing technologies. These new technologies not only improve productivity and quality of research, they also enable entirely different kinds of research,

Special thanks to the many people who helped us disseminate the survey and the individuals who took the time to complete the survey.

² Australian Government, Department of Innovation, Industry, Science and Research, *National Collaborative Research Infrastructure Strategy (NCRIS)* <<http://www.ncris.dest.gov.au/>>.

³ OAK Law Project, *Open Access to Knowledge* <<http://www.oaklaw.qut.edu.au>>.

⁴ National Collaborative Research Infrastructure Strategy (NCRIS) <http://www.dest.gov.au/sectors/research_sector/policies_issues_reviews/key_issues/ncris/>.

⁵ NCRIS Strategic Roadmap: Section 5.16 Platforms for collaboration <http://www.dest.gov.au/sectors/research_sector/policies_issues_reviews/key_issues/ncris/ February 2006>.

organisational models and collaborations across every discipline, and create new research domains that could not otherwise exist.

These capabilities serve to advance and augment, rather than replace traditional research methodologies. It is important to understand the e-Research environment to ensure that any legal framework will serve to facilitate, rather than inhibit, collaborative research and innovation.

This chapter presents a brief summary of the results from a survey conducted by QUT's Faculty of Law as part of the Legal Framework for e-Research Project.⁶

The term 'e-Research' encapsulates research activities that use a spectrum of advanced ICT capabilities and embraces new methodologies emerging from increased access to:

- broadband communications networks, research instruments and facilities, sensor networks and data repositories;
- software and infrastructure services that enable secure connectivity and interoperability;
- application tools that encompass discipline-specific tools, and interaction tools.⁷

The survey⁸ aimed to explore the nature of research collaborations and to identify common legal and project agreement problems encountered in forming research collaborations in order to form strategies to facilitate and streamline the process of e-Research in the Australian context. Specifically, the aims of the survey were to:

- identify e-Research activities and levels of engagement;
- understand the nature of the collaborative research landscape;

⁶ A full report on all of the survey results entitled *Legal and project agreement issues in collaboration and e-Research: Survey Results* is available at <<http://eprints.qut.edu.au/archive/00009112/01/9112.pdf>>.

⁷ Department of Education (DEST), *e-Research* <http://www.dest.gov.au/sectors/research_sector/policies_issues_reviews/key_issues/e_research_consult/> at 27 June 2007.

⁸ For details of the survey description and methodology see: Maree Heffernan and Nikki David, *Legal and Project Agreement Issues in Collaboration and e-Research: Survey Results* (2007) 9–13 <<http://eprints.qut.edu.au/archive/00009112/01/9112.pdf>> at 11 March 2008.

- investigate characteristics of informal collaborations and agreements; and
- explore legal issues related to data and databases.

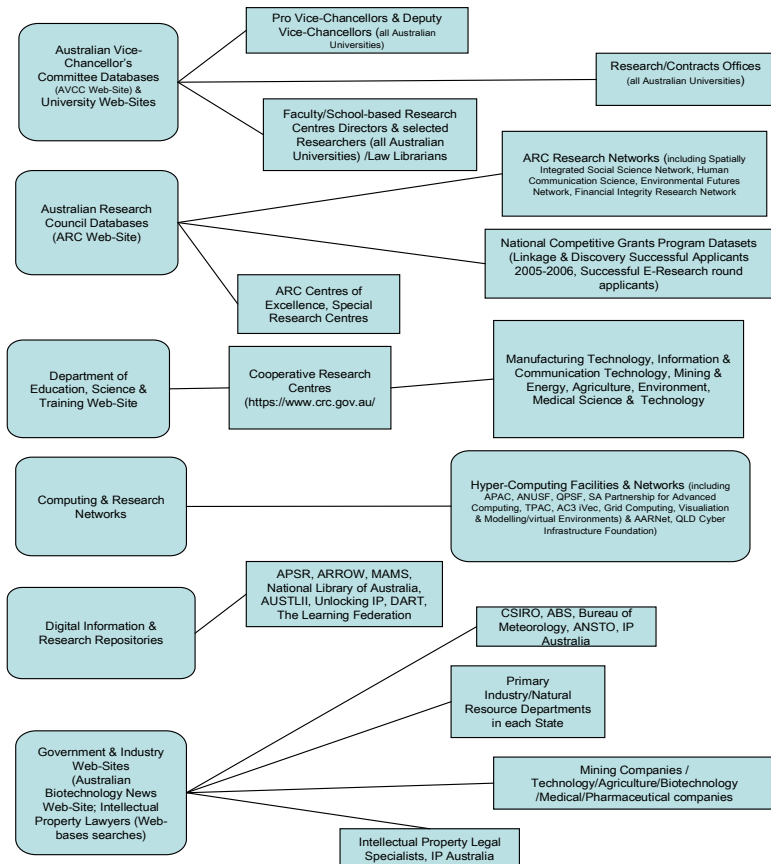
SURVEY RESULTS

Demographic Characteristics of the Sample

The questionnaire consisted of sections covering e-research (Section A), organisational/research areas (Section B), collaboration profiles (Section C), project agreement issues (Section D), databases (Section E) and data (Section F).

Of the 176 participants, 85 (or 48%) were in research roles, 66 (or 38%) were in research and/or organisational management and 25 (or 14%) were in legal or contracts roles. The majority of participants were from the University sector (64.8%), 9.1% from Industry/Commercial and 9.1% from Government sectors, 10.8% from other Research Institutes and 6.3% from law firms.

One-third of the sample stated that they are extensively involved with e-Research (only 10.3% stated that they are 'not at all' involved with e-Research). Participants were asked to describe the types of e-Research activities that their role involves. One-hundred and fifty-four participants described the kinds of activities that their e-Research involves and these were coded into broad categories based on the predominant theme of the comment. Activities described by participants included: data collection/management/modelling/visualisation and the use of databases (approximately 49% of activities); online or internet-based research (approximately 15% of activities); services to support e-Research (approximately 12% of activities); the use of communication tools (approximately 7% of activities); the dissemination of information (approximately 3% of activities); and management of e-Research activities (approximately 3% of activities).

Figure 1. Web-Based Participant Contact Sources for Legal and Research Fields

Collaboration Profile

Parties Involved in Collaborative Projects

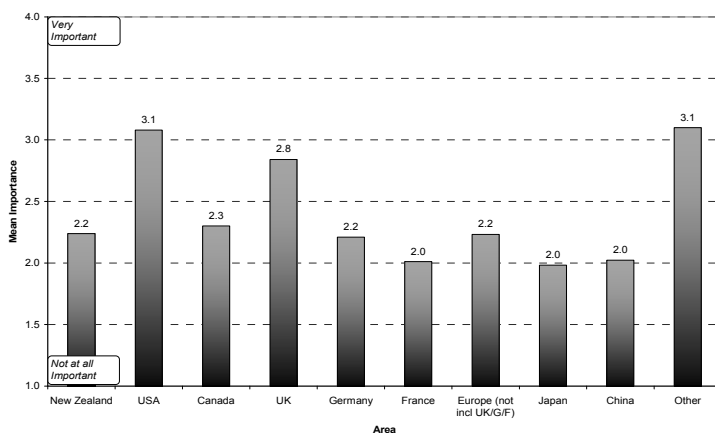
The frequency of involvement with differing parties involved in collaborative research was explored by asking respondents how frequently their collaborative projects involved industry (including commercial laboratories/R&D enterprises), universities, other research institutes, government agencies, colleagues within their organisation,

clients/customers/users, suppliers and consultants (on a scale of 1 'never' through to 4 'often').

As expected, there is a large degree of inter-university collaboration, with universities cited as the party most frequently involved in the respondents' collaborative projects, 81.3% stating that their projects often involve universities. Colleagues in their own organisation were also rated highly, with 72.2% of respondents identifying them as often being involved in their projects. Suppliers and consultants were the parties least likely to be involved in respondents' collaborative research projects. Six participants nominated parties other than those listed, such as research/postgraduate students, patent/trade mark attorneys and lobbyists.

Participants were asked to specify the most important international collaborators involved in their research projects and the results are portrayed in Figure 2. Of the specified list of countries, the USA (40% identified as 'very important'), followed by the UK (25% identified as 'very important') were identified as the most important countries to the participant's collaborative projects. Of the other countries specified by participants, India, Israel, Singapore, Thailand and islands in the Pacific were the most common.

Figure 2. Importance of International Collaborators



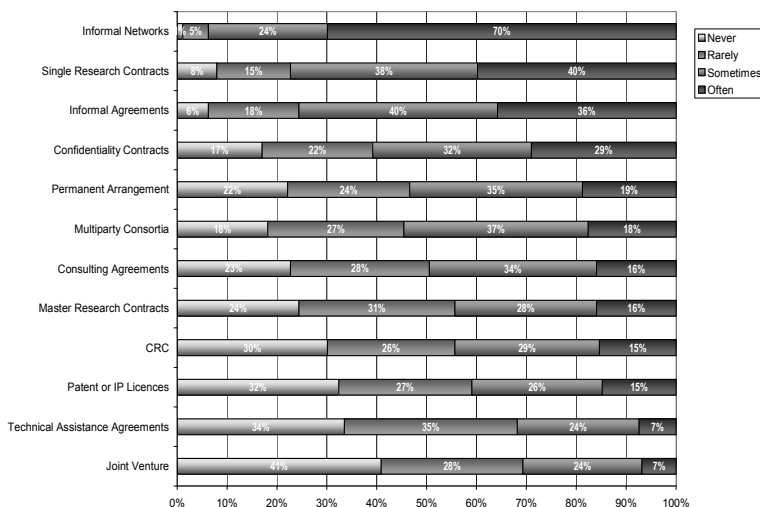
Collaborative Project Agreement Types

Participants were asked to rate the relative frequency (on a 4-point scale⁹) with which their collaborations involved a number of elements or arrangements, such as: informal networks (including informal conversations, conference interactions); informal agreements leading to co-authored publications; confidentiality/non-disclosure contracts; research contracts (for one project); master research contracts (involving multiple research projects); permanent research arrangements such as strategic alliances; multi-party research consortia; cooperative research centres; joint ventures and cross-licensing; patents/software (or other intellectual property licences); technical assistance agreements; and consulting agreements.

Figure 3 presents the relative frequency of responses to the 12 agreement/arrangement types for the total sample. 'Informal networks' (including informal conversations, conference interactions), 'informal agreements leading to co-authored publications' and 'single research contracts' were the most frequent arrangements cited. Approximately 70% of respondents stated that their collaborations often involve informal networks (including informal conversations, conference interactions), 36% stated that their collaborations often involve informal agreements leading to co-authored publications and approximately 40% stated that their collaborations often involve single project-based research contracts. Only 7% of the sample stated that their collaborations often involve joint ventures or cross-licensing (41% never) and technical assistance agreements. Approximately one-in-three participants stated that their collaborations never involve patents, software, know-how or other intellectual property licences (32.4%) or Cooperative Research Centres (30.1%). 'Commercialisation agreements' were mentioned as an additional type of agreement that is involved in collaborative projects.

⁹ 1=Never; 2=Rarely; 3=Sometimes; 4=Often.

Figure 3. Relative Frequency of Various Types of Collaboration Agreements/Arrangements



Researchers or managers (compared to those who have legal and contract roles) stated that their collaborations more often involve informal agreements leading to co-authored publications. Those who have legal and contract roles are more likely than researchers to state that their collaborations involve master research contracts or technical assistance agreements.

Managers are more likely than researchers to state that their collaborations more often involve confidentiality/non-disclosure contracts, multiparty research consortia, Cooperative Research Centres (CRC), joint ventures, patents, technical assistance or consulting agreements.

Respondents who have legal and contract roles are more likely than managers or researchers to state that their collaborations more often involve confidentiality/non-disclosure contracts, joint ventures, patents or consulting agreements.

Respondents who were from Science & Technology (compared to those from Arts & Social Sciences) stated that their collaborations more often involve master research contracts, permanent research arrangement, multiparty research consortia, or CRC.

We also wanted to gain an understanding of how informal collaborations or agreements are 'used'. Importantly, almost half of the sample stated that informal collaborations or agreements are sometimes used for detailed disclosures, and 29.5% stating that they are sometimes used to govern a whole project.

This use of informal collaboration needs to be recognised and the advantages and disadvantages need to be fully understood. Disadvantages include uncertain payoffs (barter and exchange), information gaps, credibility gaps in the information that is disclosed, risks of misappropriation and commercialisation focus (threatens the research sharing ethos).

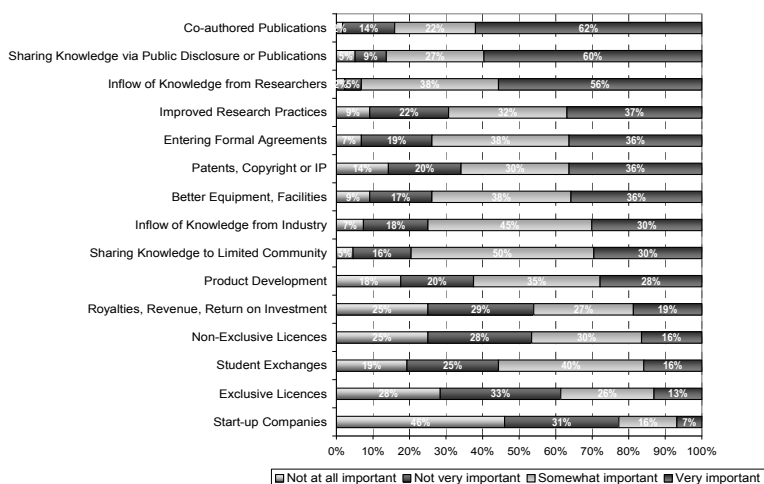
Participants were asked to specify the most important international collaborators involved in their research projects. Of the specified list of countries, the USA (40% identified as 'very important'), followed by the UK (25% identified as 'very important') were identified as the most important countries to the participant's collaborative projects. Of the other countries specified by participants, India, Israel, Singapore, Thailand and islands in the Pacific were the most common.

Collaborative Research Project Outcomes

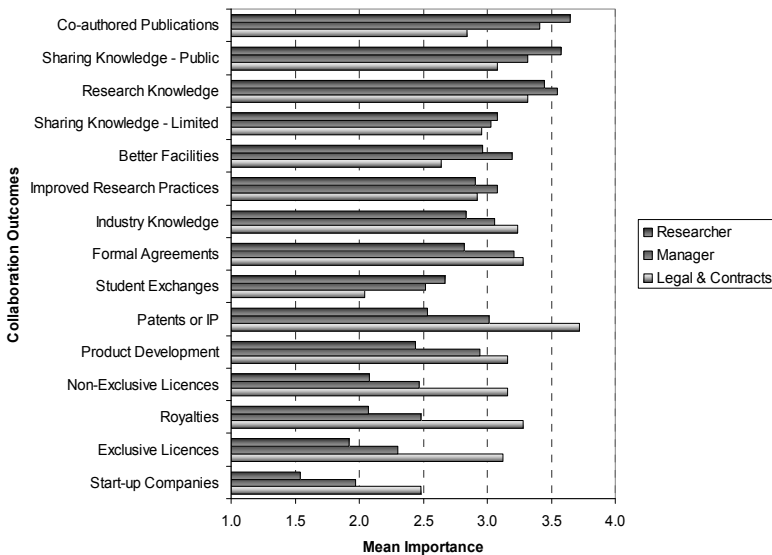
Participants were asked to rate the importance of 14 research outcomes (entering formal research agreements; patents, copyright, intellectual property; exclusive licences; non-exclusive licences; royalties, revenue, return on investment; start-up companies; co-authored publications; sharing knowledge via public disclosure or publications; sharing knowledge to limited community; student exchanges; product development, or solutions for industry/market; inflow of knowledge from industry; inflow of knowledge from researchers; better equipment, facilities; and improved research practices such as better quality, cost control, scientific evaluation) to their collaborative projects (see Figure 4).

Almost two-thirds (62%) of the sample identified co-authored publications as a very important outcome of collaborative projects (mean importance=3.44); the inflow of knowledge from researchers was identified by 60% of the sample as being very important (mean importance=3.47); and sharing knowledge via public disclosure or publications was also viewed as very important by 60% of the sample (mean importance=3.41). Figure 8 displays the mean relative importance of these research outcomes by organisational role. Three participants nominated additional outcomes such as 'improved networking' and 'rewards for communicating with others'.

Figure 4. *Importance of Research Outcomes to Collaborative Projects*



An examination of the potential differences in the importance of these collaboration outcomes by organisational role, by disciplinary area, level of involvement in e-Research and organisation sector was undertaken and Figure 5 displays the results.

Figure 5. Importance of Collaboration Outcomes by Role

Entering formal research agreements were more important for managers than for researchers. Patents or IP, exclusive or non-exclusive licences, royalties, or start-up companies were more important for those who have legal and contract roles than for managers or researchers

Co-authored publications, sharing knowledge via public disclosure or publications or student exchanges were more important for researchers than those who have legal and contract roles. Product development was more important for those who have legal and contract roles than researchers. Co-authored publications or better equipment or facilities were more important for managers than those who have legal and contract roles.

Student exchanges and product development were more important for those who are from Science and Technology than those from the Arts & Social Sciences. Inflow of knowledge from researchers was viewed as more important by those that are moderately-extensively involved with e-Research compared to those not at all-slightly involved with e-Research.

Formal agreements, patents/IP, exclusive licences, royalties, start-up companies, product development, and inflow of knowledge from industry were viewed as more important outcomes by those in government/industry compared to those in the university sector. Co-authored publications and sharing knowledge via public disclosure or publications were viewed as more important by those in the university sector compared with those in government/industry.

Critical Factors in Successful Collaborations

Participants were asked to describe the critical factors in their most successful collaborations via an open-ended question (a total of 145 comments were received). Comments were thematically coded using the following descriptors:

- Synergies and Shared Goals &/or Resources;
- Relationships & Communication; and
- Business Planning & Practice (see Appendix B of the Survey Results for the complete list of comments).

A number of participant's comments referred to a variety of factors, thus for coding purposes, the first factor specified was used to classify responses. Approximately half (49.0%) of comments made predominantly reflect the importance of research synergies and shared goals and resources, with approximately 40% of comments predominantly referring to the importance of good relationships and communication.

Comments: What do you see as the critical factors in your most successful collaborations?

Synergies & Shared/Goals Resources (approximately 49% of comments)

‘For commercial research collaborations - overlapping and complementary interests, overlapping and complementary skills, business planning, clear and honest communication paths, expectations of ongoing relationships and partnerships, joint negotiation of research, precise but flexible milestones for purely curiosity driven research - complementary and overlapping interests and skills, clear delineation of responsibility, reciprocity in interaction, good communication, opportunities for formal and informal interaction, reasonable time frames + flexible deadlines’ (Research Centre Manager, University; Arts & Social Sciences)

‘They were all run as classic skunk works where the altruistic came together informally with ... synergistic interests and the sheer determination to make it work.’ (Research Manager, Government; Science & Technology)

Relationships & Communication (approximately 40% of comments)

‘Knowledge of the people involved, the informality of the processes, goodwill between collaborators, reputations of the participants and recognition of the research outcomes likely to be achieved.’ (Researcher, Government, Science & Technology)

Project Agreement Issues

Almost one-third of the sample believe that formal agreements are always necessary (31.1%), with approximately two-thirds stating that formal agreements are sometimes necessary (68.0%). Many comments emphasised the importance of trust in collaborative arrangements. As one participant commented: ‘*If a hand shake and mutual respect won’t do it ... contracts are not going to save you from each other*’ (Research Manager, Research Institute; Arts & Social Sciences). Thirty-six participants commented on the necessity of formal agreements.

The average time taken to finalise formal collaborative research agreements (from initial contact) is 2.2 months for confidentiality/non-disclosure agreements (range 1–12 months); 3 months for simple two-party agreements (range 1–12 months); and 8 months for large, complex or multi-party agreements (range 1–30 months). As one participant commented: ‘*Legal agreements represent the largest impediment to timely research.*

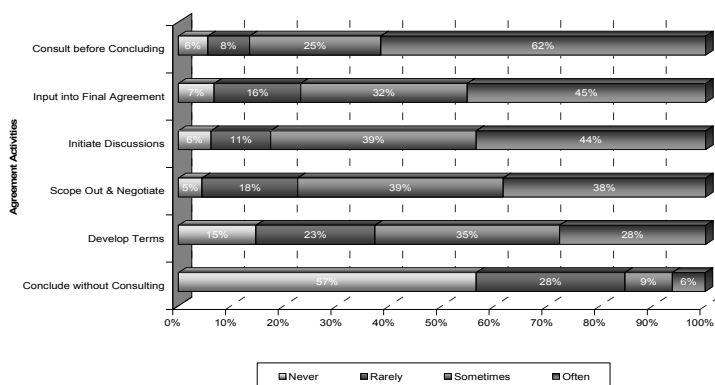
The writing of proposals and obtaining funds is the easiest and quickest part. Legal agreements require early involvement of lawyers' (Researcher, University; Science & Technology).

The majority of participants felt that they have an understanding of what the terms of their formal collaborative agreements mean. There was also relatively high agreement with a statement regarding knowing about the requirements of intellectual property ownership.

A majority of the sample were satisfied with the level of input they have into formal agreements (79.7%), with only 15.3% stating that they would like more input (5.1% stated that they would like less input). For those in research roles, 22.4% stated that they would like more input into formal agreements and for research/organisational managers, 6.1% stated that they would like to have more input.

Participants were asked to identify the frequency with which a range of activities occur in the context of project agreements (see Figure 6). Almost two-thirds (62%) of the sample often consult with others such as managers and legal/contracts advisors before concluding formal agreements (5.6% never consult and 7.9% rarely consult others). Almost half of the sample stated that they often initiate discussions with other researchers for possible collaborations (44%) and 38.4% stated that they have input into the actual form of the final agreement (and 31.6% stating that they sometimes have input into the final agreement). Over half of the sample (56.5%) also stated that they never conclude formal agreements without consultation or assistance.

Figure 6. Relative Frequency of Project Agreement Activities



There were significant differences in responses by disciplinary area, whereby those in Science & Technology fields are more likely than those in Arts & Social Science fields to 'initiate discussions with other researchers for possible collaborations', 'scope out collaborative projects, negotiate milestones and outcomes and 'have input into the actual form of the final agreement'. There were also significant differences by organisational sector, whereby those from government/industry are more likely than those from the university sector to 'assist in developing terms sheets, heads of agreement or memoranda of understanding'. Participants from universities are more likely than those from government/industry to initiate discussions with other researchers for possible collaborations.

Commencing collaborative research projects prior to the signing of agreements is a relatively common practice; with 26% stating that they 'often', and 54.2% stating that they 'sometimes', commence projects before agreements are signed (only 6.8% stated that they never start projects prior to sign-off). Comments indicate pressure surrounding timelines is often the reason for the early commencement: *'Almost always, in fact. Generally, you've got a short-ish timeline, and you can't afford to wait months for the baggling to stop. If you don't start before the contract is signed, you'll won't finish on time and end up in violation of the terms of agreement'* (Researcher, University; Science & Technology) and *'The legal and contractual processes can often be much slower than the time it actually takes to complete the research'* (Contracts Officer, Research Institute; Science & Technology).

Participants were asked to share their views on the commercialisation of research. One-hundred and thirty-five participants commented on the commercialisation of research. Many of these comments reflected the view that commercialisation is an important part of the research process (approximately 30% broadly supporting commercialisation) although there can be conflicts involved. Almost one-in-four participants commented that commercialisation should play no role in research (or a limited role) or interferes with the process and/or integrity of the research.

Figure 7 depicts the relative frequency of a range of general problems potentially encountered in negotiating formal agreements. The most frequent problems encountered by participants were 'unreasonable delays in project commencement', 'difficulties with government agencies', 'difficulties with university technology transfer offices' and

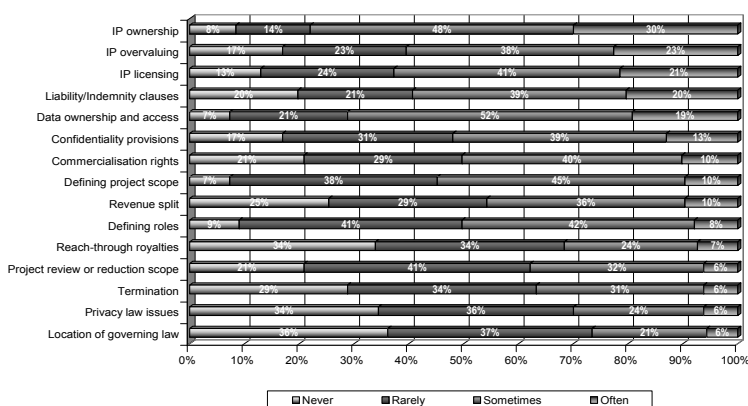
negotiation resulting in something that ‘became too complex for what the project was’. Over one-third of the sample (36%) stated that sometimes negotiation difficulties prevented the project from proceeding and that trust had been eroded.

An examination of the differences in frequencies of negotiation issues by organisational role was undertaken. Respondents who have legal and contract roles (compared to those in research roles) more often encountered the problem of the other party having all the leverage or parties having differing expectations and managers.

There were no significant differences by disciplinary area, with both those in science and technology and the arts citing unreasonable delays in project commencement as the major negotiation problem.

Participants were also asked to rate the frequency of a range of specific issues that can cause problems in negotiating formal agreements. The highest mean frequencies were attached to ‘intellectual property-ownership’, ‘data ownership and access’, ‘intellectual property-licensing’, ‘intellectual property-overvaluing it’ and ‘liability/indemnity clauses’. Half of the sample (53.1%) identified that these problems can also be a problem during the performance of the agreement.

Figure 7. *Specific Problems in Negotiating Formal Agreements*



The majority of the sample (78%) stated that when negotiating agreements they are generally able to resolve the issue of publication or

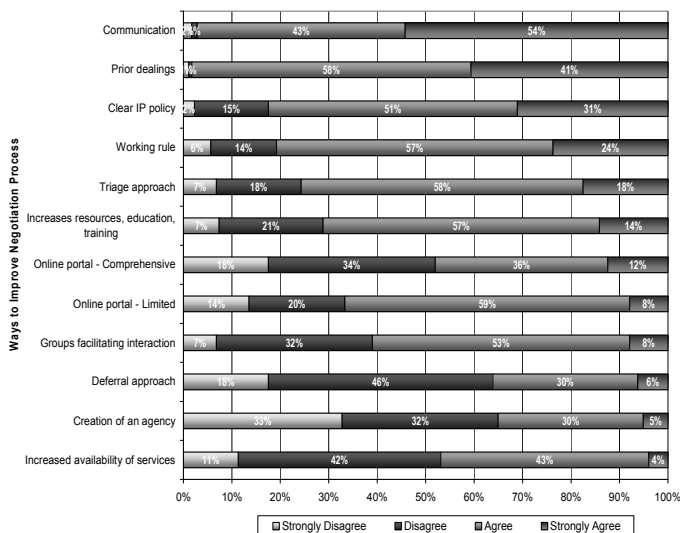
public release of results (eg by a limited delay in release to allow preservation of intellectual property rights) to their reasonable satisfaction (9% stated that it was resolved but that the delay had a serious adverse effect on their publication and 6% stated that it was resolved but there was a complete embargo on some information).

The majority of the sample had not used mediation/arbitration or court proceedings. Over half of the sample strongly agreed (16.9%) or agreed (45.2%) with the statement: 'I rely on trust to resolve disputes rather than my formal agreement'. Approximately half of the sample strongly agreed (8.5%) or agreed (45.2%) that they rely on the terms of their formal agreements to resolve disputes.

Participants were asked to rate their level of agreement with 12 statements regarding ways to improve the negotiation process (see Figure 8). The highest levels of agreement were: 'Communication, making an express effort to understand the other party's culture, objectives, drivers and mission', followed by 'Parties have had prior dealings together' and then 'Each party's organisation has a clear intellectual property policy that balances issues of access, cost recovery and return on investment. 'A generally accepted working rule that intellectual property generated in collaborative research should be divided according to relative inputs, measured by demonstrable relevance to the generated property' was rated next important then 'A triage approach, sorting agreements into those that need significant negotiation and those that do not'.

The lowest level of agreement was with the statement: 'Creating a new government agency to develop and maintain a master database of standard clauses for research contracts, issue guidelines and oversee licensing practices'.

Figure 8. *Ways to Improve the Negotiation Process*



Managers were more likely to agree than researchers that communication or increased resources, education/training for transfer offices will improve the negotiation process. Managers and researchers were more likely to agree than those who have legal and contract roles that a 'working rule' will improve the negotiation process. Respondents who have legal and contract roles are more likely to agree than researchers that increased availability of services or increased resources/education/training for transfer offices will improve the negotiation process.

Those who are from Science & Technology are more likely to agree than those from Arts & Social Sciences that a 'working rule' will improve the negotiation process. Those who are moderately-extensively involved with e-Research are more likely to agree than those that are not at all-slightly involved with e-Research that 'increased availability of services similar to contracts/technology transfer offices on a fee-for-service basis' will improve the negotiation process.

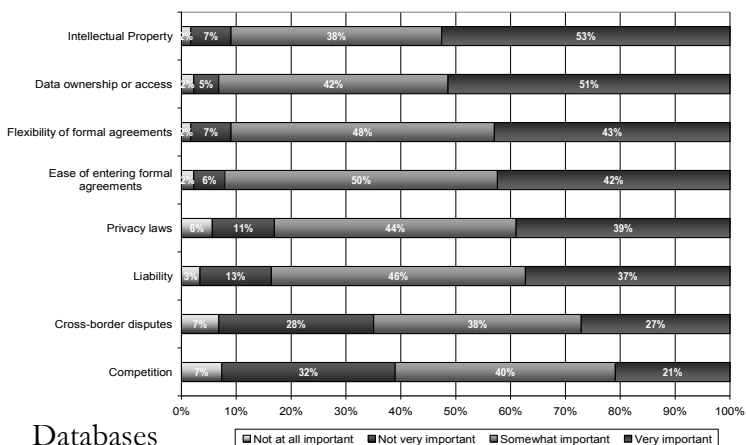
To explore views on ways to streamline documentation processes, participants were asked to rate their level of agreement with a range of

statements (see Figure 7). The statement that drew the highest level of agreement was: 'Master agreements that allow descriptions of new projects to simply be 'tacked on' are useful'. There were also high levels of agreement with the following: 'Using simple confidentiality agreements allows disclosures to occur quickly', 'Standard agreements would be customised anyway' and 'Agreements generated by assembling standard clauses would be customised anyway'. The highest levels of disagreement were attached to the statements: 'Standard agreements for different collaborations would be widely used'; 'A database of standard clauses for assembly into formal agreements would be widely used'; and 'Licensing based on the 'free/open source software' model (e.g. free access but limitations can be imposed on use, re-use, dissemination, commercialisation of content) would be widely used'. There were no significant differences by organisational role, disciplinary area or level of involvement in e-Research.

Participants were asked to rate the importance of a range of contracting issues in the context of an increase in the practice of e-Research (see Figure 9). 'Intellectual property (e.g. patents, copyright)' (53% stating that it will be 'very important' and 38% stating that it will be 'somewhat important') and 'Data ownership or access' (51% stating that it will be 'very important' and 42% stating that it will be 'somewhat important') were the issues that participants felt would become most important with the increase of e-Research. This was followed by 'Ease and speed of entering formal agreements' (42% stating that it will be 'very important' and 50% stating that it will be 'somewhat important') and 'Flexibility of formal agreements' (43% stating that it will be 'very important' and 48% stating that it will be 'somewhat important').

'Competition/anti-trust issues around research structures' was seen as the least important issue in the context of increasing e-Research activities. The only significant difference between responses by organisational role was in the view of the future importance of 'Liability' whereby those who have legal and contract roles perceive that the issue of liability will become more important with the increase of e-Research than those in researcher roles.

Figure 9. *Future Importance of Various Contracting Issues*



A total of 95 participants completed the database section of the survey (participants who selected the organisational roles of researcher and research manager). Just over three-quarters are located in universities, with one-quarter of the sample in government or industry. Approximately 37% of the 95 participants stated that they are extensively involved in e-Research (34% moderately involved; 20% slightly involved and 9% not at all involved), with 63% identifying science and technology fields and 37% identifying arts and social science fields as their area research.

Almost half (43.2%) of the sample access external databases in conjunction with their research activities on a daily basis, 36.8% on a weekly basis and 11.6% on a monthly basis. Approximately 20% are required to register for all the databases they access, whilst 22% are required to register for more than half of the databases they access, with only 20% stating that they are not required to register to access databases.

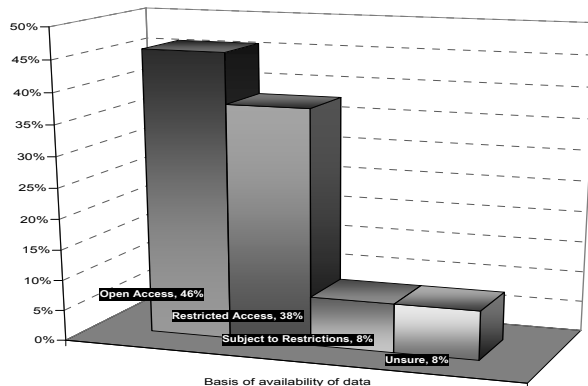
Awareness of, and compliance with, legal restrictions associated with copying, extracting or re-using information from the databases accessed was relatively high, with 74.8% stating that they have an awareness of these restrictions and 79.0% stating that they always comply with restrictive notices presented on databases. Almost half of the databases

accessed are located in Australia (47%), with over one-third located in the United States (34%).

Almost three-quarters (70.5%) of the sample felt that clearer explanations of what can be legally copied, extracted or re-used from particular databases would help facilitate their research. Fifty participants (or 52.6% of the sample) produce data or datasets that are deposited into a database. Of these participants, two-thirds (66%) created the database themselves (or their organisation created the database), and the remainder deposit into a database created by another body or institution. In terms of the location of this database, 30% are located outside of Australia. The majority of data generated is made available for access and use by other researchers (88% stating this is the case).

For those that deposit data or datasets into a database, 46% stated that it is on the basis of 'open access', whereby data is freely accessible with no restrictions on the use that can be made of it; 38% stated that it is on the basis of restricted access (such as to specific individuals or groups); and 8% stated that it is on the basis that it is subject to restrictions on the uses that can be made of the data (see Figure 10).

Figure 10. Basis of Availability of Data to Other Researchers



Many of the participants stated that their organisation does not have a policy setting out the basis on which research data should be deposited into databases for access by other researchers (53.7% stating that their

organisation does not have a policy). Table 1 presents information related to organisational policies by disciplinary area, extent of involvement in e-Research, and organisational sector.

Table 1. Presence of Organisational Policy Regarding Depositing of Data by Sector, Discipline and Extent of Involvement in e-Research

Does your organisation have a policy setting out the basis on which research data should be deposited into databases for access by other researchers?	Yes %	No %
Organisational Sector:		
University	45%	55%
Other	50%	50%
Disciplinary Area:		
Science & Technology	44%	56%
Arts & Social Sciences	51%	49%
Extent of Involvement in e-Research:		
Not at all-Slightly	36%	64%
Moderately-Extensively	51%	49%

For those participants whose organisation possesses a policy regarding the depositing of data for access by others, 84.1% stated that researchers are provided with guidelines on how the policy is to be applied in practice. Just over half (55.2%) of the 95 participants stated that they (or their organisation) prepare plans for the management and/or sharing of research data, with 62.3% of these participants stating that these plans are prepared at or around the time that grant applications for project funding are prepared. Approximately 38% stated that plans are prepared later (such as ‘during the project sometime - after analysis’ or ‘once the value of the data has been identified’).

Overwhelmingly participants felt that it would assist them to have access to a ‘plain’ English ‘how-to-guide’ explaining the legal restrictions associated with databases (89.6% stating this would assist). Of those that stated that a how-to-guide would not be of assistance, the following comments indicate potential reasons: ‘... 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' and 'it is the responsibility of the research office'.

A number of participants chose to comment on the utility of a how-to-guide: *'This may well be useful in a day-to-day sense but it would also be interesting from a digital scholarly practice perspective to see how the legal restrictions and or guidelines actually assist or impinge on scholarly practice'* (Researcher, University; Arts & Social Sciences); *'Lately we've been trying to apply creative commons licences in some cases, the availability of this licence has helped in some negotiations about data access'* (Researcher, University; Arts & Social Sciences); *'A fascinating question, given that Australia is one of the very few jurisdictions relying on copyright as the relevant property right for databases (Europe has the database right, the US does not recognise property rights in data)'* (Research Manager, University; Arts & Social Sciences).

Data

A total of 95 participants completed the data section of the survey. Almost all of the 95 participants use or generate alphanumeric data (97%), 63% use or generate images such as photographs, diagrams, graphs and/or video and 6% use or generate audio/sound data.

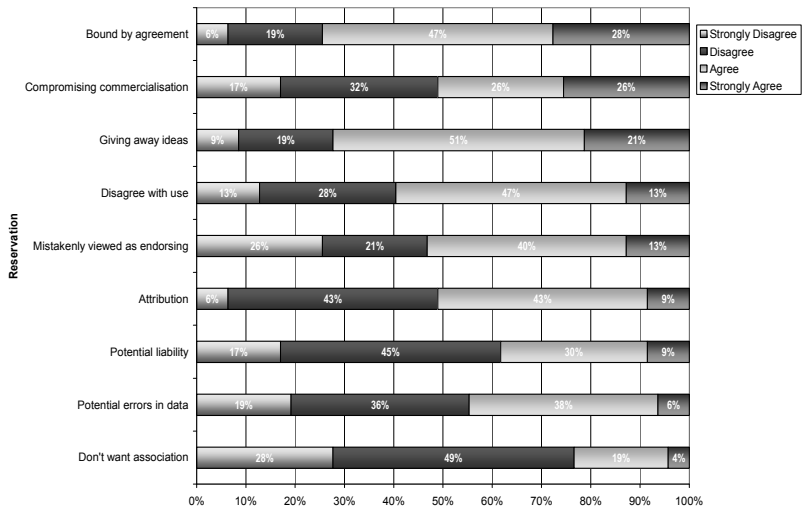
Overall, 26% strongly agreed and 63.5% agreed with the statement 'I have a clear understanding of who owns the data I use in my research projects' (10.4% disagreed or strongly disagreed). In terms of understandings of ownership of data generated, 33.3% strongly agreed and 50.0% agreed that 'I have a clear understanding of who owns the data generated in my research projects'. For those in Science & Technology fields, 39% strongly agreed and 44% agreed with this statement. In the Arts & Humanities fields, 26% strongly agreed and 60% agreed with this statement.

The majority of the sample takes steps to ensure research data is available in a form which can be readily stored and accessed (81.2%) and 56.3% stated that their organisation currently has defined mechanisms to assist in the storing and accessing of data in the long term. Comments suggest that the storage, preservation and accessing of material in the longer term can be a challenge for organisations: *'Though this is problematic ... as there does not exist the underlying infrastructure to manage this data beyond the life of the projects'* and *'Maintaining the data over the long term can be difficult as versions of software change'.*

Almost half of the 95 respondents (47 individuals or 49%) had reservations about people outside of their projects or organisation having access to their data. To explore potential reasons for these reservations, participants were asked to rate their agreement (on a scale of strongly disagree through to strongly agree) with 9 statements. Figure 11 depicts the relative agreement related to each statement. The highest level of agreement was attached to the statements: 'You are bound by a formal collaborative research agreement not to disclose data' (27.7% strongly agreeing); 'Your projects seek to commercialise the outcomes and you do not wish to compromise this', 'You do not want to give away your ideas' and 'You do not want your data to be used in research that you oppose or personally disagree with' (12.8% strongly agreeing and 46.8% agreeing).

Managers had more reservations than researchers about people outside the project or organisation having access to data because they are bound by a formal collaborative research agreement not to disclose data and reservations about not compromising the commercialisation of outcomes. Those who are from Science & Technology fields have greater reservations than those from the Arts & Social Sciences about people outside the project or organisation having access to data because the project seeks to commercialise the outcomes. Assessment of the results explored the differences in reservations by disciplinary area, organisational sector and extent of involvement in e-Research. Participants in the university sector were also less likely to agree with the statement 'Your projects seek to commercialise the outcomes and you not wish to compromise this' than participants from government and industry.

Figure 11. Reasons for Reservations Regarding External Access to Data



Almost three-quarters (74.5%) of those respondents who had concerns about people outside their project or organisation having access to data created as a result of the research project stated that their concerns would be reduced by having a legally binding agreement that clearly defined legal ownership and limited liability for the recipient's use of the data. For those that felt that such an agreement would not reduce their concerns, the following comments indicate potential reasons for this perception: *'No confidence in the law binding such people'*; *'Because ownership and liability aren't problems: ethics are'*; *'The issues of control over the use of data would not be dealt with by this'* and *'Too difficult to obtain adherence'*.

CONCLUSION

The major legal issues pertaining to establishing successful e-Research collaborations can already be broadly identified. Many of these issues are similar to those encountered in collaborative research programs using more traditional collaboration modes. However, the e-Research legal framework is potentially more complex. Collaborations by e-Research will add further complexities, which need to be identified and understood to facilitate optimisation of returns to the research participants, funding agencies and to society in general. The key points

that will impact on what legal framework for e-Research will emerge include:

1. international collaboration;
2. intellectual property and data ownership;
3. the need to reduce the friction that arises in negotiations and contracting; and
4. the recognition that informality is a key ingredient and the need for the law to accommodate and promote this dynamic elements.

To achieve its overall goals the Legal Framework for e-Research Project will:

1. acquire information on e-Research collaboration through many resources including this survey;
2. consider the institutional arrangements needed for best practice e-Research contracting/team building; and
3. present legal models for e-Research that reflect the capacity of the technologies involved and can be implemented quickly, effectively and (in many instances) in an automated way.

The survey provides a valuable insight into the Australian collaborative e-Research community. If access to knowledge is a key driver enhancing social, cultural and economic development, any legal framework proposed should advance, not hinder, such sharing. Accordingly, the Legal Framework for e-Research Project will endeavour to integrate the findings of the survey into further investigations and considerations of the appropriate legal framework for e-Research. This might include:

1. development of a dynamic collaborative e-Research agreement system along the lines of that described by Professor Mark Perry, during the Legal Framework for e-Research Conference;¹⁰
2. development of a database of key intellectual property terms that can be considered by those who are seeking

¹⁰ Mark Perry, 'Technology, Contracting and e-Research' (Paper presented at Legal Framework for e-Research Conference, Gold Coast, 11–12 July 2007) <<http://www.e-research.law.qut.edu.au/conference>>.

to draft the appropriate collaborative e-Research agreement;

3. development of guidelines for Data Management Plans and a Data Management Toolkit, as suggested by Anne Fitzgerald at the Legal Framework for e-Research Conference;¹¹ and/or
4. a handbook to assist with the timely, efficient, effective and legally robust collaborative e-Research agreements.

In many ways, our work has only just begun. We value the input received from those who participated in the survey, participated in the Legal Framework for e-Research Conference¹² and will continue to provide guidance and support for our journey ahead.

¹¹ Anne Fitzgerald, 'Building the Infrastructure for Data Access and Reuse in Collaborative Research: an Analysis of the Legal Context' (Paper presented at Legal Framework for e-Research Conference, Gold Coast, 11–12 July 2007)

<http://www.e-research.law.qut.edu.au/files/conference/audio/02_Second%20Session/02_Fitzgerald_Anne.wma>.

¹² Legal Framework for e-Research Conference, Gold Coast 11–12 July 2007 <<http://www.e-research.law.qut.edu.au/conference>>.

BIOGRAPHIES

Anthony Austin

E_mail: Anthony.austin@qut.edu.au

Anthony Austin is a Research Officer for the OAK Law Project and the Legal Framework for e-Research Project at the Queensland University of Technology, Brisbane, Australia. Anthony worked as a solicitor for 10 years in private practice before joining the OAK Law Project primarily in intellectual property and commercial law.

Anthony completed his Masters of Law degree at the Queensland University of Technology in 2007 and has worked and advised on numerous OAK Law and Legal Framework for e-Research publications including The OAK Law Project and The Legal Framework for e-Research Project Report: *Building the Infrastructure for Data Access and Reuse in Collaborative Research: An Analysis of the Legal Context*, The OAK Law Project Report: *Survey on Academic Authorship, Publishing Agreements and Open Access* and The Queensland University of Technology and ARC Centre of Excellence for Creative Industries and Innovation Guide, *CCI Blog, Podcast, Vodcast and Wiki Legal Guide for Australia 2008*. In February 2008, he participated in an international roundtable on Cyberinfrastructure, Innovation, and University Policy convened by Professor Brian Kahin at the Keck Center of the National Academies in Washington D.C. He has been accepted as a presenter at the Oxford Internet Institute's e-Research conference in September 2008.

James Casey

E_mail: lawrev@hotmail.com

James Casey is the Director of Contracts and Industrial Agreements, University of Texas at San Antonio. He has worked at a variety of research and non-research intensive universities, including Northwestern University and the University of Wisconsin-Madison. A member of the

Wisconsin Bar since 1990, James also practiced law in Wisconsin and worked in local government.

James has been a member of the National Council of University Research Administrators (NCURA) since 1995. He is currently a member of the Board of Directors, co-editor of the *NCURA Newsletter*, and a member of the Professional Development Committee. His prior service to NCURA includes serving as editor of the scholarly journal *Research Management Review* and serving as a NCURA delegate to the University-Industry Partnership Project in Washington, D.C. He has written articles for the *Research Management Review* and other publications on research administration topics and has presented at NCURA and SRA conferences. He serves as an advisor to the Program in Law and Technology at the University of Dayton School of Law, Dayton, Ohio.

In addition to research administration, Dr. Casey maintains an active intellectual agenda in the areas of transportation and public planning. He has published two books through the American Public Works Association focusing on freeway planning and history in Milwaukee, Wisconsin, and has presented on these topics to a variety of non-governmental and governmental entities.

His academic credentials include: B.A., *cum laude*, University of Wisconsin-Whitewater; M.A., international affairs, Marquette University; M.P.A., urban administration, University of Dayton; and J.D., University of Dayton School of Law, where he served on the *Dayton Law Review*.

Jessica Coates

E_mail: j2.coates@qut.edu.au

Jessica Coates is the Project Manager of the Creative Commons Clinic, a program of the ARC Centre of Excellence for Creative Innovation at Queensland University of Technology. The Clinic aims to further the implementation of the international open content licensing movement, Creative Commons, through the promotion of Creative Commons research and usage in Australia.

Jessica joins the Clinic on secondment from the Commonwealth Department of Communications, Information Technology and the Arts (DCITA), where she has spent most of the last decade as a copyright and communications policy officer. At DCITA, Jessica worked primarily

in the Intellectual Property Branch, where she took a major role in the development and implementation of copyright reform, including the Digital Agenda Amendments and the Australia-United States Free Trade Agreement. Whilst with DCITA, Jessica also worked on the ABC and SBS policy with the National Broadcasting Section and on IT usage by museums with the Collections Development Branch.

Jessica has a Bachelor of Laws from the Australian National University, and is currently undertaking a Masters in e-Law with Melbourne University.

Tom Cochrane

E_mail: t.cochrane@qut.edu.au

Professor Tom Cochrane is Deputy Vice-Chancellor (Technology, Information and Learning Support) at the Queensland University of Technology. The position heads a Division which combines the services of the Libraries, Information Technology Services, Teaching and Learning Support Services, Integrated Help Services and University Printing Services in the one structure.

In his current role Professor Cochrane's external duties include Chair, Australian Libraries' Copyright Committee; Director, Australian Digital Alliance; Director, Queensland Cyber Infrastructure Foundation; and member, Publications Board of the CSIRO. He also chairs the Australian e-Research Infrastructure Council, as well as other cross sectoral committees.

He is co-leader of the Creative Commons project for which QUT is the institutional partner for Australia. This project, together with other open access initiatives locally based at QUT, signal a long standing commitment to access to knowledge, and to research output worldwide

Terry Cutler

E_mail: cutler@cutlerco.com.au

Dr Terry Cutler is an industry consultant and strategy advisor in the information and communications technology sector. Terry Cutler has authored numerous influential reports and papers on the new Information Economy and innovation.

He currently holds the following appointments: Director of the Commonwealth Scientific and Industrial Research Organisation (CSIRO); Chairman of Board Commercial Committee; Chairman of ACID Pty Ltd, (the Australasian CRC for Interaction Design); Member of International Advisory Panel, Multimedia Supercorridor (Malaysia), 1998 - ; Member of Innovation Economy Advisory Board, Victoria; Director of MSC Technology Centre Snd. Bhd., Malaysia; Director of Multimedia University (Universiti Telekom Sdn. Bhd.), Malaysia; Council Member of Queensland University of Technology; Director of Innovation Xchange Australia Limited.

Terry Cutler's career started with Telecom Australia, now Telstra. During the 1980's he was part of Telstra's top management team, heading up major corporate restructuring around a customer focus. As Executive Director Corporate Strategy he oversighted Telstra's handling of major regulatory change and the introduction of competition.

Terry Cutler has had a longstanding engagement with public policy. He has served on numerous Government Boards and advisory bodies. Apart from his present appointments, from 1996 to 1997 he was Chairman of Australia's Information Policy Advisory Council. When Chairman of the Industry Research and Development Board from 1996 to 1998, Terry Cutler spearheaded key initiatives in promoting venture capital and industry innovation and oversaw the licensing of the initial Innovation Investment Funds. In 1999 Terry chaired Australia's National Bandwidth Inquiry.

Terry Cutler has had a longstanding engagement with cultural institutions. He was President of the Australian Centre for the Moving Image from 2002 to 2005. He served as Chairman of the Australia Council from 2001—2002, having previously chaired its New Media Arts Board. He also has previously served as a director of Cinemedia, Film Victoria, Opera Australia, the Council of the Victorian College of the Arts, and the Library Board of Victoria.

Terry Cutler is a Fellow of the Australian Institute of Management, a Fellow of the Australian Institute of Public Administration, a Member of the Institute of Company Directors, the Market Research Society of Australia, and the Australian Society of Authors. In 2002 he was awarded an honorary doctorate by Queensland University of Technology and in 2003 was awarded Australia's Centenary Medal.

In 2008, Dr Cutler was appointed as Chair of the Review of the Australian National Innovation System.

Paul Allan David

E_mail: Pad@stanford.edu

Paul Allan David is the Professor of Economics and Senior Fellow of the Institute for Economic Policy Research at Stanford University. He is Professor Emeritus of Economics and Economic History in the University of Oxford, Emeritus Fellow of All Souls College, Oxford and currently Senior Fellow of the Oxford Internet Institute. David is the author of more than 150 journal articles and contributions to edited volumes, as well as of the author and editor of several books including *Technical Choice, Innovation and Economic Growth* (1975) and *The Economic Future in Historical Perspective* (2003).

He was among the pioneering practitioners of the "new economic history," and is known internationally for wide-ranging contributions in the fields of American economic history, economic and historical demography, and the economics of science and technology. Investigation of the conditions that give rise to 'path dependence' -- the persisting influence of historical events in micro and macro economic phenomena -- is a recurring theme in his research. Two main areas of contemporary economic policy research have emerged in his work the past two decades: the evolution of information technology standards and network industries, and the influence of legal institutions and social norms upon the funding and conduct of scientific research in the public sector, and the interactions between that latter and private sector R&D. David currently leads an international research project on the organization, performance and viability of free and open source software.

Many professional honours have been bestowed upon David in the course of his career, including election as Fellow of the International Econometrics Society (1975), Pitt Professor of American History and Institutions in the University of Cambridge, as Fellow of the American Academy of Arts and Sciences (1979), Vice-President, and President of the Economic History Association (1988–89), as Marshall Lecturer in the University of Cambridge (1992), Ordinary Fellow of the British Academy (1995), Member of Council of the Royal Economics Society

(1996–2002), and Member of the American Philosophical Society (2003). He was made Professor of Economics and Economic History by University of Oxford, ‘in recognition of distinction’ (1997) and awarded a Doctorate *Honoris Causa* by the University of Torino (2003).

David’s extensive service as a consultant to international organizations has included work for the World Bank, the United Nations Commission on Trade and Development, the United Nations University Institute, the OECD, several directorates of the European Commission of the EU, the European Committee for Future Accelerators, the Economic and Social Research Council (U.K.), the Treasury and the Ministry of Science and Technology of New Zealand, and the German Monopolies Commission. He also has had extensive service-experience as a consultant to U.S. government agencies and foundations, including the National Academy of Science (National Research Council), the National Science Foundation, and the Departments of Commerce, and of Energy; the Rockefeller Foundation, the Sloan Foundation and other public and non-profit organizations. He has been a non-executive director of La Compagnie de Saint-Gobain since 2002, and recently was appointed to the board of directors of Science Commons, a not-for-profit organization founded by Creative Commons in 2005.

Claire T. Driscoll

E-mail: cdriscol@mail.nih.gov

Claire Driscoll has served as the Director of the National Human Genome Research Institute (NHGRI)’s Technology Transfer Office since 2002. At the U.S. National Institutes of Health (NIH), NHGRI is the focal point for research into the genetics of human disease. Ms. Driscoll is responsible for the overall oversight of the Institute’s intramural patent and licensing portfolio as well as the negotiation of a wide array of collaborative research agreements. She also advises staff on technology transfer policy and related matters. Prior to joining NHGRI Claire was a Senior Technology Development Manager with the NIH’s National Institute of Allergy and Infectious Diseases (NIAID). Earlier in her career she conducted research in a National Institute of Child Health and Development (NICHD) laboratory at NIH. Education: B.S. and M.S. degrees in Biology from the University of Notre Dame.

Ms. Driscoll has given presentations, primarily on biomedical technology transfer and related intellectual property and licensing topics, at various conferences including Biotechnology Industry Organization (BIO)-, Association of University Technology Managers (AUTM)- and Licensing Executive Society (LES)-sponsored events. Currently, Claire serves on the advisory boards of PXE International, a lay patient research advocacy organization, and the Genetic Alliance's Biobank, a centralized blood and tissue repository.

Erin Driscoll

E_mail: erin.driscoll@nsfa.gov.au

Erin has significant expertise in copyright law and policy, and has previously worked for the Attorney-General's Department, the National Museum of Australia and the Australia Film Commission. She was also a member of the secretariat supporting the Copyright Law review Committee in its inquiry into the relationship between copyright and contract in 2002.

Since 2003 Erin has also chaired the Copyright in Cultural Institutions Group (CICI), which is a working group of copyright and intellectual property managers across Australia, with membership including the Australian Digital Alliance, Australian Film Commission, National Museum of Australia, National Archives of Australia and National Gallery of Australia.

Anne Fitzgerald

E_mail: am.fitzgerald@qut.edu.au

Anne is a professor of law at QUT and recognised expert in the field of intellectual property law, which is demonstrated through her outstanding contributions to research, publication, training, teaching and professional practice. Anne has specialised for the past 15 years in intellectual property law, in particular its application to information technology. During this time she has gained extensive practical experience in intellectual property and technology contracting and recent hands on experience negotiating, drafting and advising on information technology and biotechnology contracts for the Queensland Government.

Anne has conducted extensive research in these fields, resulting in the publication of several books, numerous articles and book chapters on intellectual property law (particularly as it applies to digital technologies), and electronic commerce law. Since 1991, she has taught courses in the areas of intellectual property and e-commerce law to students in law, biotechnology, information technology, multimedia and electronic commerce courses, as well as to information technology professionals, writers and designers. Each year since 2003 she has taught the Intellectual Property Law course offered by Macquarie University's School of Law as a summer intensive and since 2004, she has been the co-coordinator (with John Stonier) of the Patents and Commercialisation course in the Master of Laws program at QUT Law School. Anne teaches in the undergraduate Internet Law and E-commerce Law and Technology Contracts courses offered at QUT Law School and has taught in the Cyberlaw course in Southern Cross University's summer law school program since 1998.

Anne was an initiator of the *Going Digital* series of seminars on legal aspects of e-commerce, multimedia and the Internet which were held in Brisbane, Melbourne and Hobart in 1997 and 1998 in association with QANTM Australia Co-operative Multimedia Centre. The project culminated in the publication of *Going Digital: Legal Issues for Electronic Commerce, Multimedia and the Internet* (Prospect Media, now LexisNexis/Butterworths) in August 1998. A second, completely revised, edition of the book, *Going Digital 2000: Legal issues for e-commerce, software and the internet* was published in February 2000.

Brian Fitzgerald

E_mail: bf.fitzgerald@qut.edu.au

Professor Brian Fitzgerald studied law at the Queensland University of Technology graduating as University Medallist in Law and holds postgraduate degrees in law from Oxford University and Harvard University.

He is a well-known Intellectual Property and Information Technology/Internet lawyer who has pioneered the teaching of Internet/Cyber Law in Australia. He has published articles on Intellectual Property and Internet Law in Australia, the United States, Europe, Nepal, India, Canada and Japan and his latest (co-authored) books are *Cyberlaw: Cases*

and Materials on the Internet, Digital Intellectual Property and E Commerce (2002); *Jurisdiction and the Internet* (2004) and *Intellectual Property in Principle* (2004) and *Internet and Ecommerce Law* (2007).

Brian is a Chief Investigator and Program Leader for Law in the ARC Centre of Excellence on Creative Industries and Innovation and Project Leader for the DEST funded Open Access to Knowledge Law Project (OAK Law) Project looking at legal protocols for open access to the Australian research sector and the DEST funded Legal Framework for e-Research examining the legal framework needed to enhance e-Research. He is also a Program Leader for CRC Spatial Information.

From 1998–2002 Brian was Head of the School of Law and Justice at Southern Cross University in New South Wales, Australia and from January 2002 –January 2007 was Head of the School of Law at QUT in Brisbane. He is currently a specialist Research Professor in Intellectual Property and Innovation at QUT. Brian is also a Barrister of the High Court of Australia.

Fred Friend

E_mail: f.friend@ucl.ac.uk

Fred Friend grew up by the sea in Dover, read most of the books in his local public library, and with the help of supportive parents went to study history at Kings College London. He had the good fortune to enter academic libraries at a time of growth. His first post was in Manchester University Library as a SCONUL Trainee and then as Assistant Librarian. Fred worked at several universities in the UK and obtained his first library director post at the University of Essex. This was followed by a move to University College London, where he was library director for 15 years before transferring into a role as Honorary Director Scholarly Communication, which enables him to explore new developments in information services. Fred is involved in many initiatives through work for UK and international organizations, his primary role being as Scholarly Communication Consultant to JISC. He is one of the authors of the Budapest Open Access Initiative.

Chris L. Greer**E_mail:** greer@nitrd.gov

Dr. Greer received his PhD degree in biochemistry from the University of California, Berkeley and did his postdoctoral work at CalTech. Dr. Greer was a member of the faculty at the University of California at Irvine in the Department of Biological Chemistry for approximately 18 years where his research on gene expression pathways was supported by grants from the National Science Foundation, National Institutes of Health, and the American Heart Association. During that time, he was founding Executive Officer of the RNA Society, an international professional organization with more than 700 members from 21 countries worldwide.

Dr. Greer has been a rotator and, more recently a member of the permanent staff at the National Science Foundation. He is currently Senior Advisor for Digital Data in the Office of Cyberinfrastructure. Previously, he has served as Program Director in the Division of Molecular and Cellular Biosciences, the Division of Biological Infrastructure, and the Emerging Frontiers Division. Dr. Greer recently served as Executive Secretary for the Long-lived Digital Data Collections Activities of the National Science Board and is currently Co-Chair of the Interagency Working Group on Digital Data of the National Science and Technology Council's Committee on Science.

Andrew Hayne**E_mail:** AndrewHayne@privacy.gov.au

Andrew Hayne is Deputy Director, Policy, in the federal Office of the Privacy Commissioner. Andrew leads the Office's health privacy and smartcard team, which has carriage of issues such as shared electronic health records, genetic privacy, unique healthcare identifiers and the Australian Government's proposed health and social services Access Card. A range of recent public submissions on these topics is available on the Office's website at www.privacy.gov.au. Andrew has been with the Office since 2003, and has a background in regulatory and policy positions, as well as in management academe and consulting. He holds degrees in business management and international law.

Maree Heffernan**E_mail:** maree.heffernan@qogr.qld.gov.au

Maree Heffernan works for Queensland Government and was formerly a research officer with the Legal Framework for e-Research Project and a Senior Research Officer with the Centre for Social Change Research at QUT. Graduating from the University of Queensland with first class honours in psychology in the late '90's, she is currently completing a PhD in the School of Humanities & Human Services at QUT. The focus of her research is the 'human element' in the uptake of new technologies, and she has broader interests in the areas of social justice and marginalisation. In the Legal Framework for e-Research Project, Maree has worked with the project team to develop an online survey regarding legal and project agreement issues in collaboration and e-Research and been responsible for the statistical analyses of survey data. Maree is also working with the OAK Law team on another online survey of Australian authors regarding publishing agreements. These projects build on a number of projects Maree has initiated, including: digital storytelling as a qualitative research methodology, collaborative research frameworks (including working with Indigenous communities) and the use of geographic information systems in the context of human service provision.

Richard Jefferson**E_mail:** r.jefferson@cambia.org

Richard Jefferson is the founder and CEO of CAMBIA-BiOS. He earned his Ph.D. at the University of Colorado, Boulder. In 1989 he joined the Food and Agriculture Organization as their first senior staff Molecular Biologist. He left the UN System in 1991 in order to establish CAMBIA as an autonomous private research and development institute. Richard was chosen as an Outstanding Social Entrepreneur by the Schwab Foundation. In December 2003 he was named by Scientific American to the List of World's 50 most influential technologists, and cited as the World Research Leader for 2003 for Economic Development. He was nominated as a finalist for Wired Magazine's Rave Awards for Scientist of the Year for 2005, and received the American Society of Plant Biologists (ASPB) 'Leadership in Science Public Service Award' in July 2005.

Scott Kiel-Chisholm**E_mail:** s.kielchisholm@qut.edu.au

Scott completed his Articles of Clerkship with Blake Dawson Waldron Lawyers in 2001, after working in the Insurance, Projects, Intellectual Property and Communications and Corporate Advisory practice groups. He then travelled to Silicon Valley, California USA and upon his return, Scott commenced work in the Litigation practice group of McInnes Wilson Lawyers, concentrating on the defence of professional indemnity claims. In 2004, Scott joined the Commercial Litigation practice group of Home Wilkinson Lowry Lawyers which provided broader litigation experience in project management contracting, retailing, construction, manufacturing and franchising. In an effort to progress a career in intellectual property law, Scott commenced work with Colavitti Lillas Lawyers before becoming Project Manager of The OAK (Open Access to Knowledge) Law Project in November 2005. In September 2006 Scott also became the Project Manager of the Legal Framework for e-Research Project and has been involved in the development of an online survey and international conference. In December 2007, Scott was awarded a Master of Laws specialising in intellectual property law from The University of Queensland.

Gaye Middleton**E_mail:** gmiddleton@uecomm.com.au

Gaye Middleton is an in-house counsel with Uecomm Operations, an Australian telecommunications company. Previously, Gaye has worked as a legal officer for Griffith University in Brisbane, where she worked closely with the University's academics, Office for Research and Office for Commercialisation. While employed at McCullough Robertson Lawyers, she has also advised IMBcom, the commercialisation company for the Institute for Molecular Bioscience at the University of Queensland, and various bioscience companies in respect of legal issues concerning the conduct of research. Gaye has co-authored two legal texts concerning the Internet and e-commerce, namely *Jurisdiction and the Internet* and *Internet and e-Commerce Law – Technology, Law and Policy*. She is also an Australian news contributor to *Bioscience Law Review* and *Computer and Telecommunications Law Review*. Gaye holds a Master of Laws, a

Bachelor of Laws (Hons Class I) and a Bachelor of Science from the University of Queensland.

Kylie Pappalardo

E_mail: k.pappalardo@qut.edu.au

Kylie Pappalardo is a research officer for the Open Access to Knowledge (OAK) Law Project, based at Queensland University of Technology (QUT) and led by Professor Brian Fitzgerald. She holds a Bachelor of Laws (Hons.) degree and a Bachelor of Creative Industries (Creative Writing) degree from QUT. She has also completed a Graduate Diploma of Legal Practice at QUT.

Kylie is a co-author, along with Professor Anne Fitzgerald, of *A Guide to Developing Open Access through Your Digital Repository* (2007, OAK Law Project) and *Building the Infrastructure for Data Access and Reuse in Collaborative Research: An Analysis of the Legal Context* (2007, OAK Law Project). She has also authored the recent publication, *Understanding Open Access in the Academic Environment: A Guide for Authors* (2008, OAK Law Project).

Kylie has given numerous conference presentations on behalf of the OAK Law Project, including the Australian Partnership for Sustainable Repositories (APSR) and the Australian Research Repositories Online to the World (ARROW) Adaptable Repository Workshop in Sydney in 2007. Kylie provides voluntary legal advice at the Arts Law Centre of Queensland (ALCQ) and has taught Creative Industries Legal Issues to undergraduate students at QUT. In 2007, Kylie was a recipient of the QUT Vice Chancellor's Awards for Excellence.

Mark Perry

E_mail: mperry@uwo.ca

Professor Mark Perry is jointly appointed to the Department of Computer Science and the Faculty of Law at The University of Western Ontario in London, Canada. He is also an Adjunct Professor of Law at QUT in Brisbane.

Professor Perry's research focuses on the nexus of science and law, and on autonomic computing system development. He has most recently published in the areas of digital rights management, copyright reform

and software licensing automation. He holds grants from the Natural Science and Engineering Research Council, IBM, and the Social Science and Humanities Research Council.

Prof. Perry is a frequent invited speaker at research-intensive colloquia and classes at universities in Australia, India, New Zealand, United Kingdom, United States and Canada. He has been interviewed by the media around the world for his ideas on technology law issues. In addition to serving on the Executive of ACM SIGCAS, he is a Faculty Fellow at IBM's Center for Advanced Studies, a Barrister and Solicitor of the Law Society of Upper Canada, and serves on many review and editorial boards.

David Ruschena

E_mail: david.ruschena@healthlegal.com.au

David is a Consultant at Health Legal, specialising in legislative interpretation and compliance. He has been practicing for six years and has advised public and private corporations on the reconciliation of competing obligations, in particular methods of ensuring that research does not contravene privacy legislation. He sits on an ethics committee at St Vincent's Hospital and was previously an Associate to a judge of the Federal Court of Australia. Prior to working as a lawyer, David conducted research for the Mental Health Research Institute in Victoria.

Peter Schröder

E_mail: Peter.Schroeder@dans.knaw.nl

Peter Schröder (* 1943 The Hague) lives in Leiden, The Netherlands and studied sociology (mental health and epidemiology as minors) at the University of Amsterdam. He worked as a journalist and critic of rock music before joining the Ministry of Education and Science as policy advisor for educational support systems in 1980. After a stay at Utrecht University managing the multidisciplinary research program 'Urban Networks' (1986–1990), he rejoined the Ministry at the Directorate of Research and Science Policy as policy advisor. As coordinator for scientific information policy he initiated several e-Science research programs for cultural heritage and humanities. He took part in various national and international committees advising on the updating the

regulatory and legal framework of data access regimes for research from 1990 on. Peter Schröder was co-chair of the CSTP/OECD group chaired by Peter Arzberger (San Diego) that published the report 'Promoting Access to Public Research Data for Scientific, Economic and Social Development' (2003, see <http://dataaccess.ucsd.edu> and see also Science, Vol.303, 1777 – 1778, 19 March 2004), the report that formed the basis of the OECD *Principles and Guidelines for Access to Research Data from Public Funding* (2007)

(<http://www.oecd.org/dataoecd/9/61/38500813.pdf>)

He now acts as an advisor at the newly established Data Archiving and Networked Services institute DANS (www.dans.knaw.nl) established jointly by the Royal Netherlands Academy of Arts and Sciences (KNAW) and the National Organisation for Scientific Research (NWO).

Michael Spence

E_mail: m.spence@vcc.usyd.edu.au

Dr Michael Spence is the Vice Chancellor and Principal of the University of Sydney, formerly serving as Head of Social Sciences Division at the University of Oxford. He is a consultant to the London law firm, Olswang. Michael has a comparative perspective on the law of intellectual property. His work includes articles and books on both intellectual property law and the law of obligations, with a critical focus on suggested ethical and economic justifications of the existing regimes. Michael's work on e-Science has been undertaken in collaboration with Professor Paul David of Stanford University and was first published as a report for to the Joint Information System Council of the UK in 2003, as *Towards Institutional Infrastructures for e-Science: The Scope of the Challenge*.

Fiona Stanley AC

E_mail: kristyl@ichr.uwa.edu.au

Professor Stanley is the Founding Director of the Telethon Institute for Child Health Research; Executive Director of the Australian Research Alliance for Children and Youth; and Professor, School of Paediatrics and Child Health at the University of Western Australia. Trained in maternal and child health epidemiology and public health, Prof. Stanley has spent her career researching the causes of major childhood illnesses

and birth defects. Her research includes the gathering and analysis of population data for epidemiological and public health research; the causes and prevention of birth defects and major neurological disorders, particularly the cerebral palsies and spina bifida; patterns of maternal and child health in Aboriginal and Caucasian populations; various ways of determining the developmental origins of health and disease; collaborations to link research, policy and practice; and strategies to enhance health and well-being in populations. She sits on the Prime Minister's Science, Engineering and Innovation Council as well as the Australian Statistics Advisory Council. For her research on behalf of Australia's children, she was named Australian of the Year in 2003 and in 2006 she was made a UNICEF Australia Ambassador for Early Childhood Development.

Dilan Thampapillai**E_mail:** Dilan.Thampapillai@vu.edu.au

Dilan Thampapillai is a Lecturer at Victorian University and is a specialist in intellectual property law. From 2005 to 2007, he worked at the Law Faculty at the Queensland University of Technology after working for three years with the Attorney-General's Department in Canberra. Dilan has a Master of Laws degree (LLM) from Cornell University where he specialised in intellectual property. Dilan has a Master of Commerce from the University of Sydney and a BA and LLB from the Australian National University. Dilan has also studied on a non-degree basis at Harvard University and at the National University of Singapore. Dilan was admitted to practise as a solicitor in the Australian Capital Territory in June 2003. Dilan is a researcher with Professor Brian Fitzgerald as part of the Open Access to Knowledge Law Project (OAKLAW). Dilan has published papers on peer to peer copyright infringement, copyright in the digital era and law and medicine.

Paul F. Uhler, J.D.**E_mail:** puhler@nas.edu

Paul F. Uhler is the director of the Office of International S&T Information Programs (ISTIP) and the U.S. National Committee for CODATA at the National Academies in Washington, DC, where he

works on information policy and management issues. He was the Associate Executive Director of the Commission on Physical Sciences, Mathematics, and Applications from 1991 to 1999. From 1985 to 1991 he was Senior Staff Officer for the Space Studies Board, where he directed projects on solar system exploration and environmental remote sensing programs for NASA. Before joining the National Academies, he was a Foreign Affairs Officer at NOAA in the Department of Commerce, where he worked on remote sensing law and policy and on intergovernmental agreements in meteorological satellite programs. Paul has been the director and editor of over 30 Academy studies and reports, as well as over 60 technical articles that he has published independently. He has been involved in numerous external consulting and *pro bono* activities, and speaks worldwide on a broad range of information policy and management topics. In 1997 he received the National Research Council's Special Achievement award for his work in this area. Current *pro bono* activities include serving as: official U.S. representative to the OECD Working Party on the Information Economy; co-chair of the Global Alliance on Promoting Access to and Applications of Scientific Data in Developing Countries project under the U.N. Global Alliance for ICT and Development; chair of the CODATA Task Group on Implementation Guidelines for the GEOSS Data Sharing Principles; and chair of the Legal Working Group of the Global Biodiversity Information Facility. Paul has a J.D. and Master's degree in International Relations from the University of San Diego (1983, 1984), and a B.A. in World History from the University of Oregon (1977).

John Unsworth

E_mail: unsworth@uiuc.edu

John Unsworth was named Dean of the Graduate School of Library and Information Science (GSLIS) at the University of Illinois in 2003, Urbana-Champaign, with appointments as Professor in GSLIS, in the department of English, and on the Library faculty. During the previous ten years, from 1993–2003, he served as the first Director of the Institute for Advanced Technology in the Humanities, and a faculty member in the English Department, at the University of Virginia. For his work at IATH, he received the 2005 Richard W. Lyman Award from the National Humanities Center. He co-chaired the national commission

that produced *Our Cultural Commonwealth*, the 2006 report on Cyberinfrastructure for Humanities and Social Science, on behalf of the American Council of Learned Societies, and he has supervised research projects across the disciplines in the humanities. He has also published widely on the topic of electronic scholarship, as well as co-directing one of nine national partnerships in the Library of Congress's National Digital Information Infrastructure Preservation Program, and securing grants from the National Endowment for the Humanities, the National Science Foundation, the Getty Grant Program, IBM, Sun, the Andrew W. Mellon Foundation, and others. His first faculty appointment was in English, at North Carolina State University, from 1989 to 1993. He attended Princeton University and Amherst College as an undergraduate, graduating from Amherst in 1981. He received a Master's degree in English from Boston University in 1982 and a Ph.D. in English from the University of Virginia in 1988. In 1990, at NCSU, he co-founded the first peer-reviewed electronic journal in the humanities, *Postmodern Culture* (now published by Johns Hopkins University Press, as part of Project Muse). He also organized, incorporated, and chaired the Text Encoding Initiative Consortium, co-chaired the Modern Language Association's Committee on Scholarly Editions, and served as President of the Association for Computers and the Humanities and later as chair of the steering committee for the Alliance of Digital Humanities Organizations, as well as serving on many other editorial and advisory boards. He was born in 1958, in Northampton, Massachusetts; in 1978, he married Margaret English, with whom he has three children: Bill, Thomas, and Eleanor. Further information is at:

<http://www3.isrl.uiuc.edu/~unsworth/>.

John Wilbanks

E_mail: wilbanks@creativecommons.org

John Wilbanks is the Vice President for Science at Creative Commons. He comes to Creative Commons from a Fellowship at the World Wide Web Consortium in Semantic Web for Life Sciences. Previously, he founded and led to acquisition Incellico, a bioinformatics company that built semantic graph networks for use in pharmaceutical research & development. Before founding Incellico, John was the first Assistant Director at the Berkman Center for Internet and Society at Harvard Law

School. His first technology work was at fonix, where he researched human-computer interface and pattern recognition. He also worked in US politics as a legislative aide to U.S. Representative Fortney (Pete) Stark and a grassroots coordinator and fundraiser for the American Physical Therapy Association.

John holds a Bachelor of Arts in Philosophy from Tulane University and studied modern letters at the Universite de Paris IV (La Sorbonne). He is a research affiliate at the MIT Computer Science and Artificial Intelligence Laboratory in the project on Math and Computation. He serves on the Advisory Board of the U.S. National Library of Medicine's PubMed Central and the International Advisory Board of the Prix Ars Electronica's Digital Communities awards.

INDEX

- Agreements, 67, 122, 129, 137,
141, 232, 311, 333, 355, 357, 358,
361, 367–371, 386, 390, 393, 452,
504, 512, 515, 525
- Articles, 13, 108, 288, 379, 536
- Australian Law Reform
Commission (ALRC), 177,
292, 410, 423
- Australian Research Council
(ARC), 168, 292
- CAMBIA, iii, 155, 156, 157, 158,
174, 437, 439–441, 449, 453–457,
459–462, 465, 467, 471, 473, 535
- Collaboration, viii, 61, 66–67, 71,
73, 76, 90, 111, 118, 122, 137,
141, 163–165, 169, 259, 263,
292, 333, 337, 351, 353–359,
363, 366–367, 373–376, 381–
383, 386, 390–395, 408, 497–
499, 501, 504, 507
- Competition, 94, 108, 515
- Contract, 178, 268, 328, 345, 492
Streamlining, 390–396
- Copyright
Copyright Act 1968, vi, 171,
267, 269–270, 317, 328,
483–484
- Creative Commons, 137, 144–
145, 149, 151–153, 185,
213, 220, 246, 250, 259,
273, 280, 282, 303, 330,
526–527, 530, 542
- Open Access, iv, vii, viii, 25,
147, 149, 152, 161, 186,
187, 206, 209–210, 214,
215, 220, 246, 264, 271,
274–276, 278, 282–284,
288–291, 293, 299, 302–
303, 306, 309, 444, 498,
525, 533, 536–537, 540
- Publisher, 285–286, 299, 302,
303, 304, 305, 310, 316, 330
- Cyberinfrastructure, 5, 7, 16–17,
23, 40–42, 45–46, 49, 57–59,
69, 72–73, 75, 189, 252, 525,
534, 542
- Data
Data linkage, 403, 405
Data Management, 165, 169,
281, 408, 409, 523
Data Management Plans, 523
Data Sharing, 154–155, 174,
214, 229, 249, 400, 541
Data linkage
Data linkage, 405
- DEEWR, iii, 371, 381, 497
- DEST, iii, vii, 163–166, 259, 287,
292–294, 298, 326, 381, 497,
499, 533
- DIISR, iii, 163
- e-Research, 1, iii, iv, v, vii, viii, 1,
2, 5, 25, 35, 55, 163–164, 170,
171, 264, 275, 350–358, 363,
385–386, 389–396, 400, 409,
410, 420, 423, 427, 429, 487,
497–500, 506, 507, 514–516,
518, 520–523, 525, 527, 533,
535–536
- e-Science, iv, 55–59, 60–62, 64–
67, 69–74, 76, 80, 89, 91–94,

- 96–98, 105, 107–108, 110,
118–125, 127–133, 137–138,
141–145, 221, 351, 360, 385–
387, 538
- Government, iii, vii, viii, 25, 38,
39, 63, 64, 77, 84, 121, 129,
163–165, 168, 176, 177, 195,
204–206, 216, 224, 252, 274,
277, 287, 292–293, 326, 337,
338, 340, 354, 381, 383, 384,
390, 396, 409, 414, 430, 489,
497–498, 500, 509, 528, 531,
534–535
- Innovation, iii, v, 1, 25, 28, 29,
30, 33–35, 40, 59, 61, 78, 84,
95, 126, 144, 147, 156, 163–
165, 168–169, 177, 204, 236,
265, 287, 293, 337, 350, 359,
360, 383, 384, 390, 396, 399,
407, 409, 410, 419, 452, 465,
470, 485, 497, 498, 525, 526,
528–529, 533, 540
- Institutional Frameworks, 352
- Intellectual Property, v, 1, 78, 81,
110, 122, 134, 135, 137, 147,
148, 201, 208–209, 213, 221,
223, 234, 237, 239, 249, 276,
294, 296, 312, 363, 364, 365,
373–378, 379, 457, 460, 465,
481–484, 527, 532–533, 536
- Jurisdiction, 491, 533, 536
- Lambert Model Agreements,
367, 376
- Law, v, 1, 5, 25, 95, 101, 103,
108, 122, 124, 125, 134, 137,
147, 149, 154–155, 161, 174,
176–177, 199–201, 207–209,
213, 221, 237, 240, 262, 264,
266, 268, 273, 274, 277, 279,
290, 292–294, 299, 302–304,
306, 309, 312, 321, 322, 323,
327, 328, 333, 335, 390, 406,
407, 410–414, 416, 419–421,
423, 425, 427, 432, 484, 487,
489, 491, 498, 499, 525–527,
531–533, 535–538, 540, 542
- Legal Framework, 1, iii, iv, v, viii,
ix, 1, 5, 25, 53, 147, 149, 170,
171, 264, 282, 292, 350–358,
363, 386, 389, 390–396, 410,
423, 427, 488, 492, 497, 499,
522, 523, 525, 533, 535, 536
- National Health and Medical
Research Council (NHMRC),
168, 169, 176, 177, 292, 406,
408, 415, 417, 420, 427, 430,
431
- National Institutes of Health,
223, 279, 289, 530, 534
- National Science Foundation, 5,
6, 10, 16–20, 35, 40, 41, 42, 43,
46, 57, 189, 214, 343, 530, 534,
542
- Open Access
Public Sector Information, iv,
25, 200, 205
Theses, 275, 311, 313, 314,
317, 318, 329
- Open Content Licensing, 148,
271, 280
- Open Patent Licensing, 153
- Patent, 78, 95, 122, 155, 234–
238, 292, 441, 447, 455–457,

- 460–461, 467, 473, 475, 485–486
- Policies, 1, 38, 126, 157, 200, 205, 221, 223, 224, 289, 295, 299, 352, 458
- Privacy, x, 2, 175, 176, 177, 209, 403, 405, 410–425, 427, 428, 429, 430, 486–489, 534
- ARC, 168, 169, 292, 356, 408, 525–526, 533
- Commissioner, 410–421, 423, 424, 427, 430, 489, 534
- Consent, 404, 431
- Health Information, 405, 413, 418, 431
- Health Research, 399, 402, 538, 539
- Medical Research, 291, 415, 418
- National Privacy Principles, 175–176, 414, 487, 488, 489
- Personal Information, 176, 412, 487
- Privacy Act 1988, 175, 176, 411, 418, 421, 423, 424, 427, 428, 430, 487, 488, 489
- University research, 86
- Publications, 102, 113, 173, 187, 289, 339, 483, 527
- Publishing Agreements, 302, 525
- Science Commons, iii, 144, 145, 152, 185–187, 190, 213, 220, 221, 240, 243, 246, 250–252, 280, 302, 530
- Trust, 80, 140, 141, 226, 228, 229, 260, 279, 392
- University Industry Demonstration Project (UIDP), 148