David Groenewegen
Director, Research Infrastructure
Monash University Library
Why have a repository?

- Desire to share research outputs (OS)
  - To improve research
  - To provide a public good
- Desire to share teaching and learning material (OS)
  - Open courseware
  - Learning ‘objects’
- Attack on traditional publishing models (OA)
  - Cost
  - Time to publish
  - Business model
- Because we can (OS)
  - Using new technologies
  - Improve reporting requirements
Attack on traditional publishing

- Traditional STM publishing yields outrageous profit margins
  - Elsevier: 38% = $1 billion per year
  - Springer: 34%
  - Wiley-Blackwell: 42.5%
  - American Chemical Society: $16 million

- Researchers provide content free, and also review other people’s content and edit journals for no recompense

- The publishing process can take years
How it was supposed to work

1. Set up a digital repository, preferably using open source software like ePrints or Dspace
2. Set up a self submission form for researchers
3. Promote the benefits listed two slides back
4. Wait for the pre-prints to roll in
5. Achieve critical mass
6. Journal publishers would stop selling subscriptions and move into the peer review/accreditation business
7. Access to research would be cheaper and easier
8. Job done!
The example

- arXiv (http://arxiv.org)
  - Established 1991
  - Hosted by Cornell University (previously Los Alamos Nuclear Laboratory)
  - Nearly 900,000 Physics, Mathematics, Computer Science, Quantitative Biology, Quantitative Finance and Statistics papers
  - Generally claimed to have had minimal impact on subscriptions in these fields
- Various for profit and not for profit efforts since then
International activity

- Library and discipline led repository movement
  - Concentration on pre and post print publications
    - Pre = all versions of an article up to the post print version
    - Post = The final version before the published version
    - Published version = as appears in a journal (or as a chapter, etc)
- International collaborations around Dspace, ePrints, Fedora etc
What actually happened

- **Publisher resistance**
  - Copyright as a weapon
  - Lobbying
    - Especially in US and UK

- **Researcher reactions**
  - Confusion
  - Nervousness
    - “I don’t want to get sued”
  - Apathy
    - “Most people don’t give a thought to open access”
  - Power of brands
    - “I need to publish in this journal”
  - Resistance
    - “I don’t want people to read a draft”
Result of 10+ years

- Some percentage of scholarly output is open
  - But no-one knows how much – 20%? 50%? Or how long it took for this to happen
  - Some areas better than others - in 2011, 43% of Brazilian science articles were free to read on publication, compared with, for example, 6% of US articles.
- More OA mandates than ever
- Repositories everywhere – but what are they for?
- Are they any good?
Data is next

- **G8 Science Ministers: June 12, 2013**

  i. To the greatest extent and with the fewest constraints possible publicly funded scientific research data should be open, while at the same time respecting concerns in relation to privacy, safety, security and commercial interests, whilst acknowledging the legitimate concerns of private partners.

  ii. Open scientific research data should be easily discoverable, accessible, assessable, intelligible, useable, and wherever possible interoperable to specific quality standards.

  iii. To maximise the value that can be realised from data, the mechanisms for delivering open scientific research data should be efficient and cost effective, and consistent with the potential benefits.

  iv. To ensure successful adoption by scientific communities, open scientific research data principles will need to be underpinned by an appropriate policy environment, including recognition of researchers fulfilling these principles, and appropriate digital infrastructure.

  “We decide to build on the existing work to coordinate and enable international data collaboration.”
PUBLICATIONS AND RESEARCH DATA

Research Data Collections
Journals and Newspaper Collections
Research Publications
Library Collections
Theses Collection

Content indexed by

Google Scholar - scholarly literature from many sources and indexes
OAIster - search institutional repositories around the world
Research Data Australia - discovery service for Australian research data
Economists Online - access to economics research
Trove - National Library of Australia - access to all Australian thesis and photographs
Welcome!

The key goal of The Commons is to share hidden treasures from the world’s public photography archives.

Please help make the photographs you enjoy more discoverable by adding tags and leaving comments. Your contributions and knowledge make these photos even richer.

Participating Institutions
FAQ
Rights Statement

A random sprinkling of our participating institutions...

A Commons Sampler

Search The Commons
get credit for all your research

or store it privately for FREE*

Find out more

* up to 1GB of free space

"Data which are determined to be a matter of public record and can therefore be made freely available, without restriction, are categorised as Public Access."

"Data must be properly curated throughout its life-cycle and released with the appropriate high-quality metadata."

"NERC is committed to supporting long-term environmental data management to enable continuing access to these data."
Welcome to Griffith University's Research Hub, a rich and informative guide to the University's expertise in a comprehensive array of academic fields.

The Excellence in Research for Australia (ERA) 2012 results (based on the 2005-2010 period) profile Griffith as a comprehensive, research-active university with approximately 88% of eligible researchers associated with broad (two-digit) fields of research assessed at world standard or better.
Skills needed

- People who
  - understand the technology
  - can work with researchers
  - can put in place the services and the technology that will make it look effortless.
Going forward (1)

- A view of the future. Question: What repository/storage/discovery landscape for our generated research data and outputs do we want in future?

- The overall direction of the repository has been unclear, and it is uncertain how we can tell if it has “succeeded”. Question: What is a successful repository?

- People either don’t know that it exists, or if they do, don’t know why. Question: How do we market/promote it better?

- Content has been added on an opportunistic basis, and areas of concentration have varied and been based on existing relationships, serendipity or bulk loads of Research Office material. Question: What should we focus on collecting, and how?

- The need to make data openly available in the near future to meet funder requirements. Question: How do we support Open Data?
Going forward (2)

- The existence of new services like Figshare. Question: Are there better services available, and how do we compare?
- The forthcoming RDSI data storage services. Question: Should we integrate with RDSI? If so, how?
- The increased complexity in using current software. Question: Is ours the right solution going forward?
- Research management software. Question: What do we need to do to integrate?
- The repository has a number of existing commitments and requirements (e.g. theses). Question: How do we meet these more effectively? Can we meet them all?
THANKS – QUESTIONS?

David.Groenewegen@monash.edu
@groenewegendave

THANKS – QUESTIONS?