Research Data Landscape

Laura Armstrong
Centre for eResearch

URONZ Future Horizons and the Changing Landscape of Research Management. 5-6 Sept 2019
Research Data Landscape

---

A disclaimer

Research Data – why?

Current NZ landscape

Where next?
Data Sharing and Management Snafu in 3 Short Acts

NYU Health Sciences Library
https://www.youtube.com/watch?v=N2zK3sAtr-4
Our definitions

---

**Research data** is anything collected, observed, or created for purposes of analysis to produce original research results.

**Research data management** (RDM) is the process of planning how you will collect, organise, manage, store, backup, preserve and share your data during and after your project is complete.

https://research-hub.auckland.ac.nz/#/content/53
Your experiences

In pairs

- What 1 research data thing is your organisation doing well?
- What 1 research data thing needs improving in your organisation?
Current state

PBs of data

Lacking metadata – who, what, why, when, how, and do we need to keep it?

Not finable, accessible, interoperable or reusable

Not efficient
Stakeholders

---

Researchers - integrity, collaboration, impact

Funders - value for money

Institutions - reputation, efficiency, impact

Publishers - reproducibility (& $$)

Government - policy, impact

Public - co-creators, value for money, improving lives
Funders: Data Management Plans (DMP) & Planning
Researchers and institutions

UoA Code of Conduct for Research - some RD related clauses

“Researchers should keep clear, accurate records of all research in ways that will allow verification and replication of their work by others and will evidence and support claims to and protection of intellectual property. ...

...Researchers should share data and findings openly and as promptly as possible, as soon as they have had an opportunity to establish priority and ownership claims and subject to any intellectual property requirements and contractual obligations.”

Similar across NZ.
Is this working?

https://www.otago.ac.nz/administration/policies/otago003211.html
Publisher Data Policies

1. Deposit (with DOI)
2. Citation
3. Linking or Data Availability Statement
4. Peer review of data

Elsevier
https://www.elsevier.com/authors/author-resources/research-data/data-guidelines

SpringerNature
https://www.springernature.com/it/authors/research-data-policy/journal-policies/15369670

Wiley
https://authorservices.wiley.com/author-resources/Journal-Authors/open-access/data-sharing-citation/index.html
UoA research lifecycle service tools

by Yvette Wharton
A few RDM tips
Files and folders naming

- Project/grant name/number
- Date of creation YYYYMMDD
- Initials of creator
- Description of content
- Collection method
- Version number x.y
- No spaces (no …%20..)

20170310-tmr-literature-review.docx
[date]-[creator]-[subject].[ext]

arthnz-rat-rbw-food-weights.xlsx
[project]-[animal model]-[creator]-[data type].[ext]

fr3s-140623-129C-2653-w.jpg
[studysite, depth of water]-[yyymmdde]-[tile#, treatment]-[photo#]-[photo coverage].[ext]

Retain your rights

Negotiate with the publisher

- Retaining rights to images or diagrams
  - Generic agreement
  - Elsevier
  - Bentham Science
  - Permission instead of retained rights
  - Reversion of rights

Retaining rights to images or diagrams

There are two ways in which you can possibly retain the rights to your images and diagrams:

1. Upload your images or diagrams into figshare and make them available under a Creative Commons Attribution licence. The publisher can then use them under this licence and you retain rights to adapt and re-use them.

2. Negotiate with publishers to retain the copyright for your own images and diagrams (so you can use them elsewhere) while assigning copyright for the text to the publisher. The publisher will need a non-exclusive licence to include those images and diagrams within the work to be published.
Data Publishing and Discovery Service

figshare fest NZ 2019, 30 Oct
What are your research data management tips?

- Talk to IT
- Save to shared network drive not USB or desktop!
- Consistency
- Write a data dictionary
- Plan early in the process
- Make it reusable
- Use DMPonline
- Keep it consistent within the team Use smart, user-friendly tools
- Ask around
- Version control
- Start planning early
- Contact Laura :)

Go to www.menti.com and use the code 53 09 93
Changing behaviours

---

Digital data – born digital and digitised

Future proof ethics submissions – scope and sharing (& publishing)

Use ‘University managed storage’ –

Publish (opt-out rather than opt-in)

‘Delete’ not ‘destroy’ research data – often impossible to destroy data saved to tape/long-term storage
Data ethics and AI

---

Significant funding opportunities

Open science + big data + data sharing + machine learning (+ public cloud services) = potentially re-identifiable data

What does this mean?

Possible solution - **homomorphic encryption** allows computation on encrypted data without having to decrypt the data first.

Internationally

Who has *maturity - can/should we follow their example?

University of Edinburgh - UG Data Mindfulness

University of Melbourne - all HDRs

Australian Government - ARDC, ACOLA AI, AI Ethics Framework

Tu Delft - Research Data Framework Policy

Future

---

More Open Science – co-design

Data sharing agreements – institutional processes

Data as assets – persistent identifiers across research lifecycle

Instrument data services – coming soon at UoA

National data services?

Research Data Policy – data stewardship

Machine actionable DMPs – iterative, connected to systems and actions
Which RD areas are important to your institution?

- Inst. Research Data policy & governance: 62
- Data ethics: 75
- Security: 76
- Data sharing (FAIR): 58
- Data Sovereignty (CARE): 55
- Infrastructure: 59
- RD skilled workforce: 52
- Something else: 7
A FAIR future?

Findable
Accessible
Interoperable
Reuseable
Thank you.

Laura Armstrong

Senior eResearch Engagement Specialist
Centre for eResearch, University of Auckland

E: l.armstrong@auckland.ac.nz  T: @finlibash & @eResearchUoA