The tribal approach academia takes to research data management

Keynote talk at LIBER 2017 workshop:
Research data support meets disciplines: Opportunities & challenges
Presented remotely
5 July 2017

Dr Danny Kingsley
Head of Scholarly Communication
University of Cambridge
@dannykay68
SLIDES: https://www.repository.cam.ac.uk/handle/1810/265166
The Cambridge story

• In 2015 the Office of Scholarly Communication was launched with a focus on:
  – Meeting open access policies
  – Meeting new research data sharing requirements

• On 1 May 2015 one of the funding bodies (EPSRC) said they were going to check compliance with their requirement to share data (policy in place since 2011)

http://osc.cam.ac.uk/
First - getting library staff up to speed

Introductory workshops for Supporting Researchers in the 21st century programme held in 2015. Three sessions, 103 people altogether attended.

http://osc.cam.ac.uk/training/supporting-researchers-21st-century-programme
More involved than they thought
Horses for courses

https://www.amazon.co.uk/Academic-tribes-territories-Intellectual-Disciplines/dp/0335206271
Imagine where history sits on this scale.

Image: CC-BY-NC 2.5 XKCD  https://xkcd.com/435/
Data? What data?

And many, many others:-
- Raw instrument readings
- Processed data
- Analysed data
- Genomic data
- Microscopic photos, western blot images and measurement
- Spreadsheets
- Videos
- Surveys and interviews
- Field notes
- Lab books
- Physical samples
- Protocols
- Code
- Software
Immediate researchers’ reactions:

• This is not my priority
• Why would I do that?
• People will steal my results!
• Data management is a waste of time
• Nobody will understand my data
• It would take me 5 years to find all my data!
• Telling researchers the rules had changed and they needed to share their data was not successful
• We realised we needed to start at the beginning not the end
• Setting things up well from the start means *everything* works better

Many RDM messages are agnostic.

The need for file naming conventions

Copyright: http://10pm.com/
Three years on – what are these?
File naming convention:

TILS Document Naming Convention

Document naming for the TILS Division should follow this convention:

```
GDL_TILSDocNaming_V1_20090612.docx
```

- **A prefix** shows the document type
- **The document title** describes the content
- **The version number**
- **The date** in the format yyyyymmdd

http://www.data.cam.ac.uk/files/gdl_tilsdocnaming_v1_20090612.pdf
Data organisation

Is this your desktop?
What works 2

• Blog post: “Strategies for engaging senior leadership with RDM – IDCC discussion”
  – Refer to doomsday scenarios and risks to reputations
  – Provide high profile cases of fraudulent research
  – Refer to the institutional mission statement / value statement
  – Collect horror stories of poor data management practice from your research community
  – Don’t be afraid to talk about the problems you are experiencing

https://unlockingresearch.blog.lib.cam.ac.uk/?p=1435
What would happen if?
At least 2 backups at 2 locations

Store at home!

Free software to manage backups (there is plenty of free software!):
http://www.2brightsparks.com/download-syncbackfree.html
What works 3

• Research Data Champions project
  – Engaging researchers and administrators interested in RDM
  – Sent out invitation.
  – Expected 10 responses and had 40 applications
    • Researchers – 21
    • Administrators - 4
    • Library staff – 10

https://www.data.cam.ac.uk/intro-data-champions
Who did we get?

• Primarily from the STEM sector
• A couple of Social Science representatives
• No-one from the Humanities (we are not there yet!)
• Spreading the love:
  – Rosie Higman (who started this) is now at Manchester University
  – Marta Teperek is moving to TU Delft

Conclusions

• You need to be flexible in the RDM space
• What works for one might not work for another
• Language matters! (‘FAIR’ not ‘open’)
• Be inventive and try different ways of approaching the issue
• Good Research Data Management really helps all aspects of the research endeavour
Thanks!

Dr Danny Kingsley
Head of Scholarly Communication
University of Cambridge
@dannykay68