Facilitating effective data management and sharing in University College Dublin: new library services within a campus-wide approach

Open access and research data management: Horizon 2020 and beyond
University College Cork, April 14th & 15th 2015
Outline

• Drivers
• Cycles
• Research Lifecycle Services
• Concerns: Researchers and Library Staff
• Concluding Questions
Drivers

• 2012: Creation of new Research Services Unit within UCD Library
  • Remit includes the development of specialised information services for qualitative and quantitative data

• 2012: Move of the Irish Social Science Data Archive from UCD’s Geary Institute to UCD Library

• UCD Strategic Plan: Increase the quality, quantity and impact of our research, scholarship and innovation (Objective 1)

• Changing open access landscape
Research Data Cycle

Collect
- Locate existing data
- Collect new data
- Plan consent for sharing
- Plan data management (formats, storage, security, backups etc.)
- Capture and create metadata

Process
- Enter data, digitise, transcribe, etc.
- Merge data
- Check, validate, clean data
- Anonymise where necessary
- Describe data
- Manage and store data
Research Data Cycle

Analyze
- Statistical analyses
- Data visualisation
- Prepare data for preservation

Preserve
- Migrate data to best format and suitable medium
- Backup and store data
- Create metadata and documentation
- Archive data

Disseminate
- Share data (or not)
- Establish copyright
- Promote data
Research Data Cycle Within the Research Lifecycle

Discover
- Data reviews
- Funding opportunities

Create / Analyse
- Specialist software & tools
- Data visualisation

Manage
- DMP for funding applications
- Organise & manage data
- What to consider...

Post Project
- Repository management
- On-going curation
- Audits

Disseminate / Publish
- Deposit options & requirements
- DOI; citation; licensing; metadata
“Discover” Services

- **Data reviews**

- **Data sources**
  - [www.earthchem.org/portal](http://www.earthchem.org/portal)  
    - Geochemistry data portal
  - [www.ucd.ie/issda](http://www.ucd.ie/issda)  
    - Irish social sciences / public health quantitative data
  - [http://datadryad.org/](http://datadryad.org/)  
    - Open repository for journal articles and supporting datasets in evolutionary biology

- **On what basis can data be reused?**
  - End user licenses
“Create / Analyse” Services

- Use of specialised software and tools
  - ArcGIS, QGIS
  - SPSS, PSPP
  - Nvivo
  - Omeka

Dublin Parish data downloaded from the CSO

Use of Omeka to create an online narrative
http://blacklib1969.swarthmore.edu/
“Manage” Services

Horizon 2020
Research and Innovation Actions
UCD Help Pack

One mention of the Library:
“In addition to laboratory equipment, you should also detail the IT Resources and Library resources and collections that are available”.

Need better integration into workflows, at point of information need (e.g. application, award granted, at end of project).

UCD Research:
“Research Project Lifecycle”
“Manage” Services


- Add new section:
  - 2.9 Funders’ open access requirements

- The researcher should be aware of funders’ open access requirements and mandates, both in relation to research publications and data. These are outlined at http://libguides.ucd.ie/data/funders

- Re-number 2.9 to 2.10

- Then add following boxes:

  - **UCD Library carries out the following:**
    - Supports researchers in helping them to comply with funders’ open access requirements.
    - Supports researchers in assisting with the drawing up of data management plans which may be required by funding bodies.
    - Supports researchers in the provision of metrics to enhance their funding applications.
    - Supports researchers in the provision of infrastructure, curation and other related services (e.g. metadata) for cultural heritage funding applications.
“Manage” Services

• Data management plans
  • Funder requirement to submit a data management plan with application
  • Library promotes use of checklists, templates and tools such as DMPOnline
  • Library works with Research Office = one service

• Effective data management throughout the active research phase - how to organise, structure, store, and share research data
  • Research efficiencies
  • File organisation, naming, version control etc.
  • Metadata and documentation
  • Plan consent for sharing; ethics
  • Storage, backups, security
This page gives you an overview of your plan. It tells what your plan is based on and gives an overview of the questions that you will be asked.

<table>
<thead>
<tr>
<th>Plan name</th>
<th>Minor Harbours of Southeast Ireland</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
<td>-</td>
</tr>
<tr>
<td>Grant number</td>
<td>-</td>
</tr>
<tr>
<td>Principal Investigator/Researcher</td>
<td>Elizabeth Shotton</td>
</tr>
<tr>
<td>Principal Investigator/Researcher ID</td>
<td>orcid.org/0000-0002-6703-0648</td>
</tr>
<tr>
<td>Plan data contact</td>
<td>353(1)7162703 <a href="mailto:elizabeth.shotton@ucd.ie">elizabeth.shotton@ucd.ie</a></td>
</tr>
<tr>
<td>Description</td>
<td>Pilot study to document the physical fabric of minor harbours using LiDAR scanning from which 3D models will be generated. The evolution of maritime engineering technology is illustrated in minor harbours due to their incremental development.</td>
</tr>
</tbody>
</table>

This plan is based on:

The first phase of the DCC default template

<table>
<thead>
<tr>
<th>Sections</th>
<th>Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Collection</td>
<td>- What data will you collect or create?</td>
</tr>
<tr>
<td></td>
<td>- How will the data be collected or created?</td>
</tr>
</tbody>
</table>
What data will you collect or create?

3D scans using LiDAR (Light Detection and Ranging) technology will be taken for 3-5 minor harbours. 3D models will be generated using scan data to enable wider access. Photographs and detailed notes on indicative construction techniques will be taken at the same time.

Survey of existing records and/or historical accounts will be undertaken for each harbour to supplement existing inventory, Ports, Piers and Harbours of Ireland (2 volumes) held by National Monuments Archive and running data collected from fishermen by the Underwater Archaeological Unit of the DACG.

A comprehensive record will be created for each harbour which will include:

1. LiDAR scan data (original data set, but difficult to access)
2. 3D model (3Ds format for ease of access)
3. Photographic record
4. List of historical research
5. List of historical records (based on Ports, Piers and Harbours of Ireland)
6. Images of historical drawings if available
7. Commentary on historical development including notes on engineers, owners, techniques used, current condition)

This data will serve to underpin research into the evolution of maritime
How will you manage any ethical issues?

There are no ethical issues involved in the capture of the data. The management of access to the data will be discussed with the owners of the structures and other information sources and letters of agreement outlining terms of access obtained and stored with the data set for each harbour.
“Manage” Services

That NACDA ensures that when informed consent is being sought for the collection of personal data, that the consent request clearly specifies consent for data sharing beyond the original research.

“Data archiving is a process, not an end state where data is simply turned over to a repository at the conclusion of a study. Rather, data archiving should begin early in a project ....archiving [becomes] part of the research method”. (Jacobs and Humphrey, 2004 “Preserving Research Data.”)
“Manage” Services

• What data?
• What do you want to do with these items?
  • Page through a volume like a book; search for text; view like an online exhibit; add to existing digital collection; long-term storage and curation...
• What metadata do you have?
• What are the rights / permissions of the items?
• What funding options are available to help support any aspect of the project?
• Decade of anniversaries opportunities
“Disseminate / Publish” Services

- Which data repository/ archive?
  - What are their requirements?
- Visibility, retrievability & citeability aspects
  - DOIs (e.g. DataCite - B.L., EZID)
  - ORCIDs
  - Data citation
- Does everything need to be shared?

Earthchem metadata creation tool:
www.earthchem.org/data/templates
“Post Project” Services

- Repositories
  - Ongoing curation
  - Collection building
    - Policies e.g. collection development
    - Relationship-building with depositors
- Services
  - Help with follow-up queries
  - Good practice guides
  - Data collections’ publicity
  - Data curation and profile raising that facilitates data reuse by others
    - GUI Register of Use
  - Services to depositing organisations e.g. usage statistics
  - Relationship-building
“Post Project” Services

- Working directly with Schools to audit their data collections
  - School of Archaeology wants to sort out their “data mess”
  - Will raise questions but will also assist in understanding their data management practices and procedures; and their needs and requirements
- Data Asset Framework from the DCC: identify, locate, describe and assess research data assets
Research Data Management LibGuide
http://libguides.ucd.ie/data

• Research Data Management - a Cross-Campus Service

“The purpose of this Library guide is to bring together University resources and services to facilitate researchers in the production of high quality data with potential for long-term use.”

Library as coordinator of scattered and possibly invisible services
  • Library as “neutral”
  • Create one service despite the variety of service providers
  • Relationship-building / management with key stakeholders
  • Management of collaborative activity and advocacy…who does what….

• UCD Office of Research Ethics
• UCD IT Services (storage and security)
• UCD Innovation (intellectual property, licensing and commercialisation)
• UCD Research (funding bids)
• UCD Library (data management plans, metadata, deposit to a relevant archive, etc.)
Researchers’ Concerns

• Acknowledge and address
  • Professional concerns e.g. commercialisation, patents
  • Legal e.g. use of third party data

• I don’t have time for all of this
  • Efficiencies, data loss prevention
  • Factor data management into the workflow
    • Easier and more cost-effective

• Where exactly can I deposit my data?
Data Sharing Messages

• Continually changing landscape
  • Funders
  • Journal publishers
  • Institutions

• Evidence of citation advantage
  • Piwowar HA, Vision TJ. (2013) Data reuse and the open data citation advantage. *PeerJ* 1:e175 [https://dx.doi.org/10.7717/peerj.175](https://dx.doi.org/10.7717/peerj.175)

• Competitive advantage in funding applications

• Open data, data mining...benefits to the wider research community
    • “text or datasets are crawled by software that recognizes entities, relationships, and action - helps researchers draw new conclusions among disparate data and is emerging as an important area of scholarly research”

• Metadata and documentation
• DOIs
  • De-mystify
  • Aid others being able to make sense of your data
  • Facilitate visibility, retrievability and citeability
What would you do if you lost your research data tomorrow?

“Research Data Management isn't principally about complying with policy - at heart it means helping you to complete your research, share the results, and allow you to get credit for what you have done.”

• Professor Kevin Schurer, Pro-Vice Chancellor (Research and Enterprise), University of Leicester
http://www2.le.ac.uk/services/research-data

• Quoted in SCONUL’s Research Data Management: Briefing for Library Directors (March 2015)
http://www.sconul.ac.uk/sites/default/files/documents/SCONUL%20RDM%20briefing.pdf
Library Staff Concerns:
“I don’t know enough about this area”

- Training kits
  - [http://libguides.ucd.ie/data/tutorials](http://libguides.ucd.ie/data/tutorials)

- Digital Curation Centre (JISC supported)
  - [www.dcc.ac.uk](http://www.dcc.ac.uk)
    - How-to guides, case studies, training and policy advice

- Research Data Alliance
  - [https://www.rd-alliance.org/](https://www.rd-alliance.org/)
    - ‘Good and bad data stories

- JISCMail RESEARCH-DATAMAN
  - [https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=RESEARCH-DATAMAN](https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=RESEARCH-DATAMAN)
    - Very active and useful list

- Reports
  - JISC: Directions for research data management in UK universities (March 2015)
Concluding Questions

• How can a seamless RDM service make the life of a researcher easier?
• How do we effectively collaborate with others to provide a seamless service?
• How do we demonstrate good RDM with benefits such as career progression?
• How can we embed good data management practices in research workflows?
• How do we “sell” our key messages?
• How do we gain the support of key Campus stakeholders?
• How do we provide infrastructure and dedicated research data storage and curation?
• Do we need a University policy?
• How do we develop discipline-specific training and consultancy programmes and services?
• How do we develop our own skills?
• What specialist skills will be needed and how do we provide them?
• What specialist spaces and equipment will be needed?
• What RDM-related costs can be included in a funding application?
• How will we fund all of this?
“Follow up. We're the Munsters anyway. “Come on boys”. As the men followed on they were mowed down. He had the flag tied on to the bayonet. Got right over the German parapet. When inside he waved the flag and the comrade rushed on and in. The flag soon fell and he never returned.

“I got this book today (31st Jan. 1915). Indeed I require it badly, as there are all so many little things one requires to write and remember”

“The bullets rattled and cracked all over the old distillery, and the boys sat quietly – chatting, smoking and enjoying themselves round the fires before their departure for the trenches.”
Open access and research data management: Horizon 2020 and beyond

This event was funded by FOSTER through the European Union’s Seventh Framework Programme http://www.fosteropenscience.eu and organised by

• University College Cork http://www.ucc.ie

• Teagasc http://www.teagasc.ie

• Repository Network Ireland http://rni.wikispaces.com