Welcome to the RDARI Session @ RDA 18!

“Challenges in improving interoperability within your research institution”

1. Introduction to RDARI and scene setting - 5 mins
2. Challenges at UCL designing an interoperable infrastructure (James Wilson) – 15 mins
3. Challenges at the University of Cape Town designing an interoperable infrastructure (Renate Meyer, Kimi Keith) – 15 mins
4. Plenary discussion on drivers and challenges – 45 mins
5. Wrap up and Next steps – 10 mins
Research Data Architectures for Research Institutions

“The Research Data Architectures in Research Institutions Interest Group is primarily concerned with technical architectures for managing research data within universities and other multi-disciplinary research institutions. It provides insight into the approaches being taken to the development and operation of such architectures and their success or otherwise in enabling good practice.”
Research Data Interoperability Challenges at UCL

James A J Wilson

November 2021
The Research Data / Information Lifecycles

Planning

New Ideas

Funding bid?

Data collection

Data Storage

Data Analysis

Deposit & Data Publication

Discovery & Access

Review & Open Access Checks

Researcher Profiles

Long-term Preservation

Measure Impact

REF Funding (UK)

• UCL Research Data Repository (for data)
• UCL Discovery (for papers / monographs)
• UCL Explore
• Bibliographic / citations databases
• Specialist data repositories
• Commercial and/or generalist data repositories

• DMP Online

• Worktribe
  • Je-S (UKRI submissions)
  • Various funding body grant-trackers
  n.b. Not all research projects are funded by external funders, some are simply undertaken informally, covered by salary

Experiments

Clinical trials

Images

Literature

Sensors

Datasets

Software Models

Surveys & Interviews

+ metadata

Ideally using RDSS or Data Safe Haven, but may include:

• Laptop hard drives
• Departmental Storage
• UCL Filestore
• OneDrive
• UCL RSpace
• DropBox or other commercial cloud services
• Storage managed by collaborators
• B2Store (EU service)
• Many others!

• UCL Research Data Repository (for data)
• UCL Research Publication Service (for papers / monographs)
• Research Fish

• Worktribe
  • Je-S (UKRI submissions)
  • Various funding body grant-trackers
  n.b. Not all research projects are funded by external funders, some are simply undertaken informally, covered by salary

• UCL RSpace
  • Huge variety of systems

• IRIS (drawing from various sources)
• TBC

• Manual checks in RPS and RDR by Library Open Access and Research Data Management Teams

• UCL Research Data Repository (for data)
• UCL Research Publication Service (for papers / monographs)
• Specialist data repositories
• Commercial and/or generalist data repositories

• Myriad, Kathleen – RITS HPC facilities
• National / International HPC facilities
• Locally-installed software (a vast number of specialist software tools are used for data analysis)
Metadata interoperability

- Some systems record project-based metadata
- Some systems record dataset-level metadata
- Some systems record whatever user-defined metadata the researchers care to add
- Most don’t follow recognized standards
- No two systems capture exactly the same fields
Integrating the Research Data Storage Service with HR and Figshare
Architectural Principles

• Point-to-point integrations
• Single vendor (!)
• Data lake / Extract Transform Load
What problem(s) are we trying to fix?

• Replication / Reproducibility
• Confident re-use of data
• Some disciplines have tackled problems at disciplinary level, but a lot haven’t