Preserving Data Across the Research Lifecycle

Integrating Rosetta with Figshare.

Angus Taggart - Library Systems Manager
Jez Cope - Research Data Manager
Research Data

So what is it?
The evidence gathered by the researcher to support their findings.
For the researcher to define.
citable (it has a DOI).

Requirement of funders to make it available.
Research Lifecycle

- Planning
- Active Research
- Publishing
- Preservation
- Reuse
- Discoverability & Impact
Research Lifecycle

Research Services

Planning

Reuse

Active Research

Preservation

Discoverability & Impact

Publishing
Research Lifecycle

- Research Services
  - Planning
  - Reuse
- Active Research
- Preservation
- Publishing
- Discoverability & Impact
- ORDA
Research Data

So what does it look like?

2017_06_13.whl  white_hole_model.dat  rel_time_dilation.zip
Research Data in Rosetta

Unidentified Format

- 79%
- 21%

Identified

Initial tests
The Dream

- Ingest all Sheffield records from ORDA.

Todo:
The Dream

- Ingest all Sheffield records from ORDA.
- Ingest all DOI versions.

Todo:
The Dream

- Ingest all Sheffield records from ORDA.
- Ingest all latest DOI versions.
- Utilise existing OAI-PMH functionality.
The Dream:

- Ingest all public datasets from ORDA
- Ingest latest DOI versions.
- Utilise existing OAI-PMH functionality.
- Use Rosetta’s HTTPS transport to collect files

Todo:

- XSL to filter null datasets.
- Include required information in OAI
The Dream:

- Ingest public datasets from ORDA
- Ingest latest DOI versions.
- Utilise OAI-PMH functionality.
- Use Rosettas HTTPS transport to collect files
- Display notification flag on ORDA

Todo:

- XSL to filter null datasets.
- Include required information in OAI
- Make api endpoint for flag on ORDA
- Make a flag on ORDA ui
ORDA

/or2/oai?metadataPrefix=mets

mets
FigshareID
Versioned DOI
File URI

XSLT

Rosetta

Rosetta

Rosetta

Rosetta

Collection of Files

Metadata Only
Embargoed
Today

- Testing Figshare Implementation
- Figshare API callbacks planned in v5.4
- ORDA launch - 29th June
- Launch Integration at Sheffield
Tomorrow

- Ingest of Embargoed Content
- Format Identification and Guidance