Making more of your research data

Gary Brewerton
(Middleware and Library Systems Manager)
What is research data?
Your research data: What is it?

• “Data that is collected, observed, or created, for purposes of analysis to produce original research results” – Boston University

• “The contents of the researcher’s hard disk which I remove and store in a locked cabinet after they’ve left just in case someone asks for it” – anonymous IT manager
Your research data: Why is anyone interested in it?

- Underpins and validates research findings
- May contain confidential (e.g. intellectual property) or sensitive information
- Useful to other researchers
- Funders and publishers may require it to be made available
- Promotes researchers (e.g. citations) and institutions
Your research data: What are the challenges?

• Unlike traditional research outputs such as articles and reports which are mostly text, research data can take a variety formats (e.g. spreadsheets, images, video, computer code, non-digital artefacts)
  • Varies dramatically in size (i.e. storage requirements)
  • May only exist in a propriety format
  • Will the format of the data be readable in ten years?
What it meant for us
Your research data: Institutional drivers

- Meet the “open data” mandate of research funders, in particular as it relates to metadata discovery and archiving of research data
- Further promote Loughborough research and attract future collaborators
- Provide a streamlined and intuitive workflow for research staff
Your research data: What we initially did

• In 2013 formed a Steering Committee consisting of active researchers (across various disciplines) and staff from the research office, library and IT services
  • Engage with local stakeholders and funders
  • Survey Loughborough researchers about their data requirements
  • Produce draft policy on research data management
  • Evaluate possible solutions and recommend a way forward
• Loughborough did have a current research information system (CRIS) and institutional repository (IR) but neither was considered suitable for research data
  • Existing deposit process didn’t cater for research data
  • Storage requirements for research data were uncertain but known to be greater than currently available
  • Existing solutions not considered appropriate for promoting research data
Your research data: Identified gap in provision
Your research data: Give us the money!

- Having identified the gap in our provision we asked for funding to develop and operate a new service
  - Create a new platform (interface for our researchers and others)
  - Provide adequate storage
  - Additional staffing to manage service: Research Data Manager and IT support (0.25 FTE)
- And got it!
Implementing a solution
Your research data: Choosing a solution

• Steering committee had already evaluated the market and recommended an innovative solution asking two suppliers to work together
  • Arkivum: provider of large scale, long term, safe digital archiving
  • figshare: cloud-based service allowing researchers to upload and share datasets, media, papers, presentations, etc.
Your research data: Our goal

- Loughborough Researcher
  - Symplectic Elements (CRIS)
    - Publications and associated metadata
    - DSpace (IR)
      - Publications
  - Authentication
  - Single sign-on (IdP)
    - Research data and associated metadata
      - Authorisation

- figshare
  - Data archiving
  - Research data
    - Public

- Arkivum
  - Research data
Your research data: Implementing the solution

- Kick off meeting with suppliers in September 2014
- Used Basecamp for project management/discussions
- Fortnightly catch-up calls using Google Hangouts
- Face-to-face meetings every two months (supported by the provision of lemon drizzle cake)
- Went live with solution at the end of April 2015
Where are we now
Your research data: Where are we now?

- **Soft launch of service**
  - Advocacy campaign to inform researchers about what’s happening (including email notifications, hands-on workshops, departmental presentations, meetings with research groups and one-to-one sessions)
  - Initially operated as a mediated service by the Library
  - Continuing to gathering feedback from users to further develop the service
Your research data: Depositing data

• Easily add new content

• Drag and drop file(s) to deposit

• Minimal data entry required (just 4 fields)

• Option to embargo items

• All secure archived for 10+ years
Your research data: Discovering data

- Various way of navigating content
- Visualisation of data
- Drag and drop citations
- DOIs when published
- Tags for finding related content
Your research data: Developments since launch

- Using *figshare for institutions* as a data source for *Symplectic Elements* (our CRIS)

- Projects: collaborative archive for local researchers

- Collections: groupings of items

- Author profiles
Your research data: Where we are

- Loughborough Researcher
  - Symplectic Elements (CRIS)
    - Publications and associated metadata
  - Single sign-on (IdP)
    - Authentication
    - Harvested metadata
    - Research data and associated metadata
    - Authorisation
    - Data archiving
  - figshare
    - Research data
    - Data archiving
  - Arkivum
  - Public
    - Publications
What does the future hold
Your research data: Future developments

- Further development of the solution
  - Approval process to allow Library to “sanity check” submissions
  - Allow non-Loughborough staff to collaborate on projects
- Further integration with Symplectic Elements
- And possibly migrating the content of our IR to the new platform
Conclusions
Your research data: Lessons learnt

• What worked well
  • The work of the Steering Committee allowed us to rapidly adopt a solution and get ahead of the game
  • Collaborative development with commercial partners has been a really positive experience
  • Keeping metadata requirements to a minimum aids uptake

• What would we do differently
  • Begin the advocacy earlier so that we could now demonstrate the solution to researchers without having to explain why
Your research data: 65 Thanks

Arkivum Matthew Addis • Cathy Brode • Andy Conway • Mark Ellis • Chris Pates • Jeremy Smith
Nik Stanbridge • Jo Windel DataCite Rachel Kotarski • Elizabeth Newbold DCC Alex Ball • Sarah
Jones • Liz Lyon • Jonathon Runs • Angus Whyte Digital Science Stephen Crawley • Alan
Hyndman Figshare Florin Apetrei • Livia Ciobanu • Daniel Gavrilla • Christopher George • Mark
Hahnel • Daniel Hook • Becky Westbrook Project Management Board Jeff Brown • Sally Brown
Malcolm Cook • Fidelma Hannah • Bal Sandhu • Danielle Sharp • Richard Taylor • Dave Temple
Peter Townsend Project Working Group Katie Appleton • Izabelle Bimson • Gareth Cole • Angela
Crawford • Louise Denniff • Naomi Dungworth • Marcus Enoch • Gary Page • Bethan Whitelaw
Steering Committee John Feather • Elizabeth Gadd • Martin Hamilton • Julian Highfield • Matthew
Inglis • Melanie King • Susan Manuel • Adrienne Muir • Myra Nimmo • Zoe Stockdale • Ron
Summers • Jamie Timmons • Rob Wilby • Clare Wydell • Helen Young Symplectic Sabih Ali • Nick
Anderson • Daniel Enright • Tom Letcher Others Stephen Ashurst • Hannah Baldwin • Jason
Cooper • Jonathan Walters
THE END

Any questions?