Jean Sykes

Developing a shared national research data service for UK HE: feasibility and costs

Keynote Item

Original citation:
This version available at: http://eprints.lse.ac.uk/25595/

Available in LSE Research Online: November 2009

© 2008 Jean sykes

LSE has developed LSE Research Online so that users may access research output of the School. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LSE Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain. You may freely distribute the URL (http://eprints.lse.ac.uk) of the LSE Research Online website.
Developing a shared national research data service for UK HE: feasibility and costs

Jean Sykes
13 October 2008
SHARE conference
What’s the problem?

• Research data – an untapped resource
• Often unstructured and inaccessible
• Lack of coherent policies and standards
• Many formats and disciplines
• HE Library and IT services under pressure to help (mainly storage) - unsustainable
What’s the problem?

• Whole data lifecycle, not just storage
• Creation, selection, ingest, storage, metadata, retrieval, review, preservation
• Access/analyse/synthesise others’ data
• It’s the management of the data that needs a UK-wide approach
Aims of the project

• Develop an understanding of the UK’s current and future research data service needs
• Test the feasibility of a UK-wide shared service for the management of research data
• Develop a detailed business plan
• Avoid reinventing the wheel in any proposed solution
The approach

- SERCO appointed as consultants
- Governance from Steering Committee and Project Management Board
- Over 40 Stakeholder institutions identified and representatives interviewed
- Iterative process to achieve maximum understanding
The approach

- Four case study universities: Bristol, Leeds, Leicester, Oxford
- Questionnaires and focus groups at first three
- Complementary internal project at Oxford dovetailing with UKRDS
- Total number of individuals consulted: 700
What have we learned?

• Over 360% growth in data volume expected over the next 3 years
• c50% of data estimated to have a useful life of up to 10 years; 26% seen as having indefinite retention value
• Most research is currently held locally
What have we learned?

- 21% use a national or international facility
- Most share data within research teams (12% do not); 18% share via a data centre
- 43% would like access to others’ data
- Those with no access to a national facility are particularly keen on a UKRDS
What have we learned?

• There’s a lot of good stuff out there already:
  • Data repositories with considerable skills and resources (eg UKDA, NERC, ADS)
  • Data Curation Centre Life Cycle Model
  • Data management plans (Wellcome)
  • JANET and JISC’s integrated infrastructure
What have we learned?

- More good stuff out there:
- JISC’s Federated Access Management scheme to provide trusted access to data
- Recent studies by JISC (eg data audit framework, data handling skills shortage)
- Recent studies by RIN (eg data stewardship, preservation costs)
What have we learned?

• Much work is going on in this area in Europe (particularly Germany), USA, Canada and Australia

• A service is probably feasible that brings:
  - coherence to current fragmented and incomplete provision
  - a framework for future development

• A large centralised service is unlikely to work
Are we in step?

- Cliff Lynch in Nature: “Researchers need to be obliged to manage their data with as much professionalism as they devote to their experiments” and:
- “Universities and funding agencies need to provide and support curation facilities, tools and training”
Are we in step?

- Canadian government says there is “urgent need for action to propel Canada into a new and transformational data-intensive paradigm for Canadian research” and:

- “The research process generates huge amounts of data that are an important part of Canada’s scholarly record and hold enormous potential as an additional discovery and problem-solving tool for researchers”
Are we in step?

- Alliance of German Science Organisations wants to equip scientists and scholars “with the infrastructure best suited to meeting their research needs” and:
  - “In the digital age, this entails digital access to publications [and] primary research data, available to the user without costs or other barriers”
- This is in the explicit context of “increasing Germany’s competitiveness in world-class research”
Interim report recommendations

- UKRDS should add coherence to what is already there and should be a service which:
  - Is governed by a set of policies
  - Delivers a set of processes
  - Supports a set of standards
  - Is measured by an agreed set of KPIs
  - Is funded to ensure it can be relied on long term
Activities / processes provided

Communities of interest
The next steps

- Cost existing provision
- Estimate likely costs of filling gaps and developing provision without a UKRDS
- Develop process models
- Develop a value proposition
- Develop a business plan and a governance model
The next steps

- Design a service portfolio, organisation and legal structure and find it a home, ideally in an existing UK HE entity
- Engage stakeholders in the detailed proposals
- Present draft report to Steering Cttee 31 Oct
- Produce final report for HEFCE December 2008
The next steps

• Seek interim funding in 2009 for a start-up service in co-operation with case study institutions and some existing providers
• Hold an international conference in February 2009 to promote the proposed service and highlight developments in the UK and abroad
Key messages

• The study is looking for the optimum business case to address the sustainability of what researchers need
• It’s not just about storage
• It’s about the management of the whole data lifecycle
• It will not reinvent wheels
Key messages

• Many building blocks already in place in UK
• A UKRDS would embrace rather than replace existing facilities
• There are also significant gaps to be filled
• It’s about the leverage of more research value and a higher global research reputation for the UK
Conclusion

• A UKRDS is likely to be feasible
• Need to work out a practical business case and organisation to make it happen
• Need support of funders, HEIs and other stakeholders to get it established and stable for the long term