

# The UK Research Data Service feasibility study

Martin Lewis

Director of Library Services & University Librarian  
UKRDS Project Management Board



# What's the problem?

- research data – an underexploited resource
- often unstructured and inaccessible
- lack of policies and standards
- many formats and disciplines
- HE library and IT services under pressure to help (mainly storage) - unsustainable

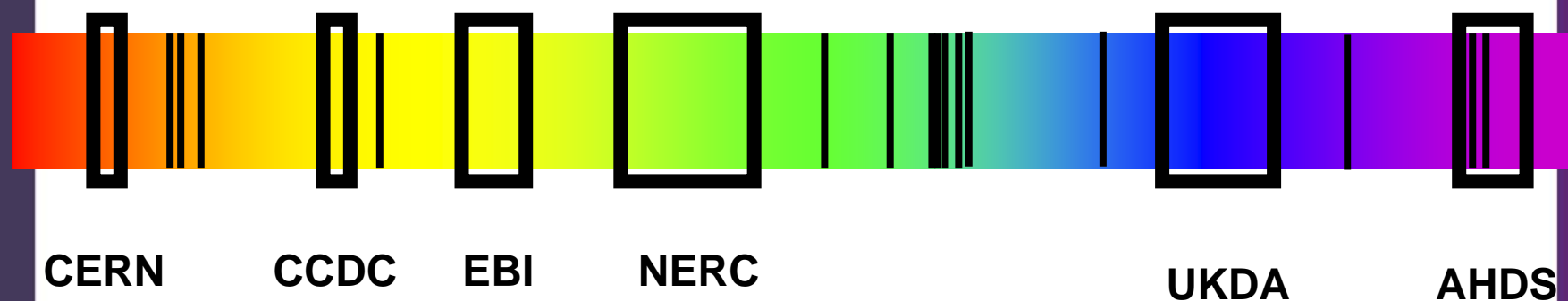


# What's the problem?

- whole data lifecycle, not just storage
- creation, selection, ingest, storage, metadata, retrieval, review, preservation
- access/analyse/annotate/synthesise other researchers' data
- significant gaps in current provision



# Research data curation spectrum

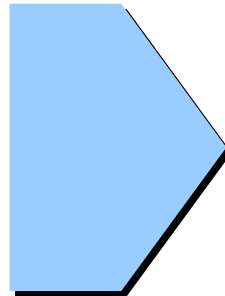


# UKRDS origins

 HM TREASURY  dti department for  
education and skills

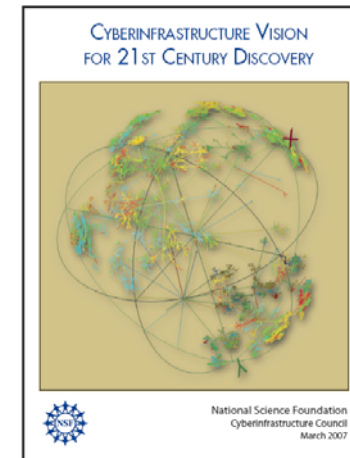
Science & innovation  
investment framework  
2004 - 2014

July 2004



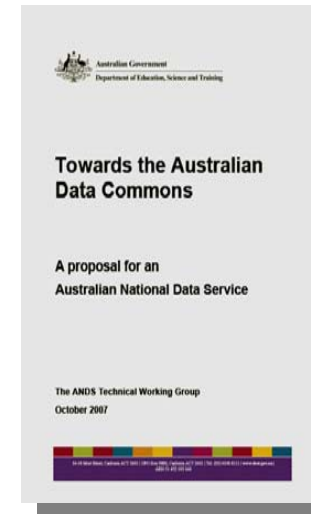
# Meanwhile...

- NSF Cyberinfrastructure vision for 21st century discovery
- Initial budget allocation \$127M
- US Interagency Working Group on Digital Data established by NSTC
- NSF Sustainable Digital Data Preservation and Access Network Partners ([DataNets](#))



# Meanwhile...

- Australian National Data Service proposal
- ESFRI (European Strategy Forum on Research Infrastructures) road map and green paper



# Is the UK losing its lead?

- eScience Core Programme a world leader
- but no funding for data/information management in first phase
- no shortage of reports....
- [Dealing with data](#) (Lyon, 2007)
- now we need services not projects



# Building blocks (1)

- some of the infrastructure in place now: JANET, JISC IE, NGS, OMII-UK, NaCTeM etc
- Digital Curation Centre has scoped many of the technical challenges
- existing good practice in data management in some Research Councils (ESRC, NERC)
- business model for some national HE shared services well established eg JANET



# Building blocks (2)

- extensive JISC project work: data preservation costs, data audit framework, skills/roles/career development of data scientists, federated tools and services, linking data with published outputs
- current RIN studies: publication and quality assurance of research data; data stewardship principles



# HEFCE shared services

- Spring 2007 HEFCE shared services call for collaborative feasibility studies
- Research Libraries UK and RUGIT defined scope with support from SCONUL, UCISA and JISC
- working group established to draft the bid
- led by LSE



# HEFCE shared services

- high level of HEFCE interest
- support from other stakeholders including HEFCW, SFC, DENI
- RLUK/RUGIT group developed full ITT for a major consultancy
- HEFCE approved £200k budget
- further contributions from JISC, RLUK and RUGIT



# Partners and stakeholders

- 40+ stakeholder agencies including HE Funding Councils and Research Councils, RCUK, BL and TNA, UUK and Russell Group, SCONUL and UCISA
- 4 case studies: Bristol, Leeds, Leicester and Oxford universities



# Aims of the project

- develop understanding of UK's current and future research data service needs
- identify priorities for action in consultation with key stakeholders
- develop options for the shared service from “do nothing” to a managed national service
- develop a detailed business plan



# Aims of the project

- include assessment of costs and benefits in options appraisal
- indicate scale of investment required and estimate return on Investment
- present outline governance proposals
- produce accessible, high quality report
- **keep research data management on the national agenda**



# Progress so far

- consultants appointed March: SERCO
- Project manager appointed: John Milner
- Project definition completed and work packages and timescales agreed
- Meetings with case study participants in April/May
- Steering Committee chaired by Prof John Wood (Imperial)



# Progress so far

- visits and telephone interviews used in four case study universities
- questionnaires distributed to research groups
- key stakeholder interviews ongoing
- individuals consulted so far: over 700
- draft final report discussed at Steering Cttee  
31 October



# What have researchers said?

- over 360% growth in data volume anticipated over the next 3 years
- c50% of data is estimated to have a useful life of under 10 years; 26% seen as having indefinite retention value
- most research data is held locally



# What have researchers said?

- 21% use a national or international facility
- Most share data within research teams; 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



# 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

- some important building blocks are already in place
- a UKRDS would embrace rather than replace existing facilities
- there are significant gaps to be filled
- it's about the leverage of more research value and a higher global research reputation for the UK



# UKRDS in a wider context

- the UK is well positioned to benefit from
  - early momentum on e-science
  - leadership on institutional repositories - for online pre- and post-print research publications
  - excellent network and access management infrastructure
- building the national data management infrastructure is now a priority



# More information

- UKRDS website
  - [www.ukrds.ac.uk](http://www.ukrds.ac.uk)
- UKRDS project manager
  - John Milner ([j.milner@lse.ac.uk](mailto:j.milner@lse.ac.uk))
- UKRDS conference
  - February 26, 2009, Royal Society

