

Blue Ribbon Task Force on sustainable digital preservation and access — a view from the UKDA

Dr Matthew Woollard

**Director-Designate,
Head of Digital Preservation and Systems,
UK Data Archive,
University of Essex**

**Wellcome Collection Conference Centre
London, 6 May 2010**



Overview

- Introduction to UK Data Archive
- Report in General
- Agenda for Research Data
 - Priorities for selection, curation and retention
 - Preservation mandates
 - “Data under stewardship”
 - Cost reduction
 - Third-party agreements
- Other issues

Introduction to UK Data Archive

- An “umbrella” organisation;
- Provides infrastructure and shared services for:
 - Economic and Social Data Service;
 - census.ac.uk;
 - Secure Data Service;
 - RELU-DSS;
 - History Data Service.
- Existing customer base;
- Continuous research and development;
- Standards compliant (various levels);
- 40 years experience;
- Legal Place of Deposit since 2005.

Stakeholders I

- Repository Funders (mainly ESRC/JISC)
 - Access to funded research outputs
 - Access to other data allowing research
 - Generally clear in terms of rewards and incentives

Stakeholders II

- Data Creators / Owners
 - Researchers
 - Verifiability / Validation
 - Good sharing culture
 - Visibility
 - Generally clear in terms of rewards and incentives
 - Government
 - Open Access (publicly funded research)
 - Accountability
 - Data Protection / Privacy
 - Legal Deposit
 - All
 - Security from data loss

Stakeholders III

- Data Users
 - Core users from social sciences in Higher Education
 - Others from wider research environment (university, government, public bodies.)
- Growth of customer base foreseeable
 - Public
 - Democratization of data
 - Future users

Sustainable economics for a digital planet: Ensuring long-term access to digital information

- Poses three important questions
 - What do we preserve?
 - Who is responsible for it?
 - Who pays?
- Content types
- Organisational actions
- Recommendations

Priorities for selection, curation and retention

Files will be of higher value in 30 and/or in 75 years than now	:	8
Files will be of equal value now, in 30 and in 75 years	:	27
Files will be of equal value now and in 30 years, but lower in 75 years	:	14
Files of highest value now, of lower value in 30 and 75 years	:	22
Files of highest value now, lower in 30 years, lower still or nil in 75 years	:	8
Files of immediate value only	:	30

Or:

Equal or higher value in the future	45%
Lower (or no) value in the future	55%

Source: *PRO Keeper's Report, 1975*

Priorities for selection, curation and retention

- **UKDA**
 - Discipline coverage
 - Geographic coverage
 - Temporal coverage
 - Time series and panel data
 - Thematic coverage
 - User request
 - ESRC funded or ESRC investment
 - Meets technical criteria and ingest/preservation/access costs are external to core funding
 - Legal and ethical considerations

Priorities for selection, curation and retention

- **ESDS Collections Development Committee**
 - Increase research productivity
 - Stimulate new networks/collaborations
 - Transfer knowledge to industry
 - Improve the national skills base in data handling
 - Increase productivity
- Nothing yet on long-term interest

Preservation mandates

- ESRC Data Policy (2000)
- ESRC Research Data Policy (2010, forthcoming)
 - OECD Principles
 - Clear roles and responsibilities
 - Focus on ESRC-funded research data (at present)
- UKDA/ONS agreement

“Data under stewardship”

- ESRC Research Data Policy (2010, forthcoming)
- Responsibilities for Information Collection
 - Impact for stakeholders
 - Repository strategic planning

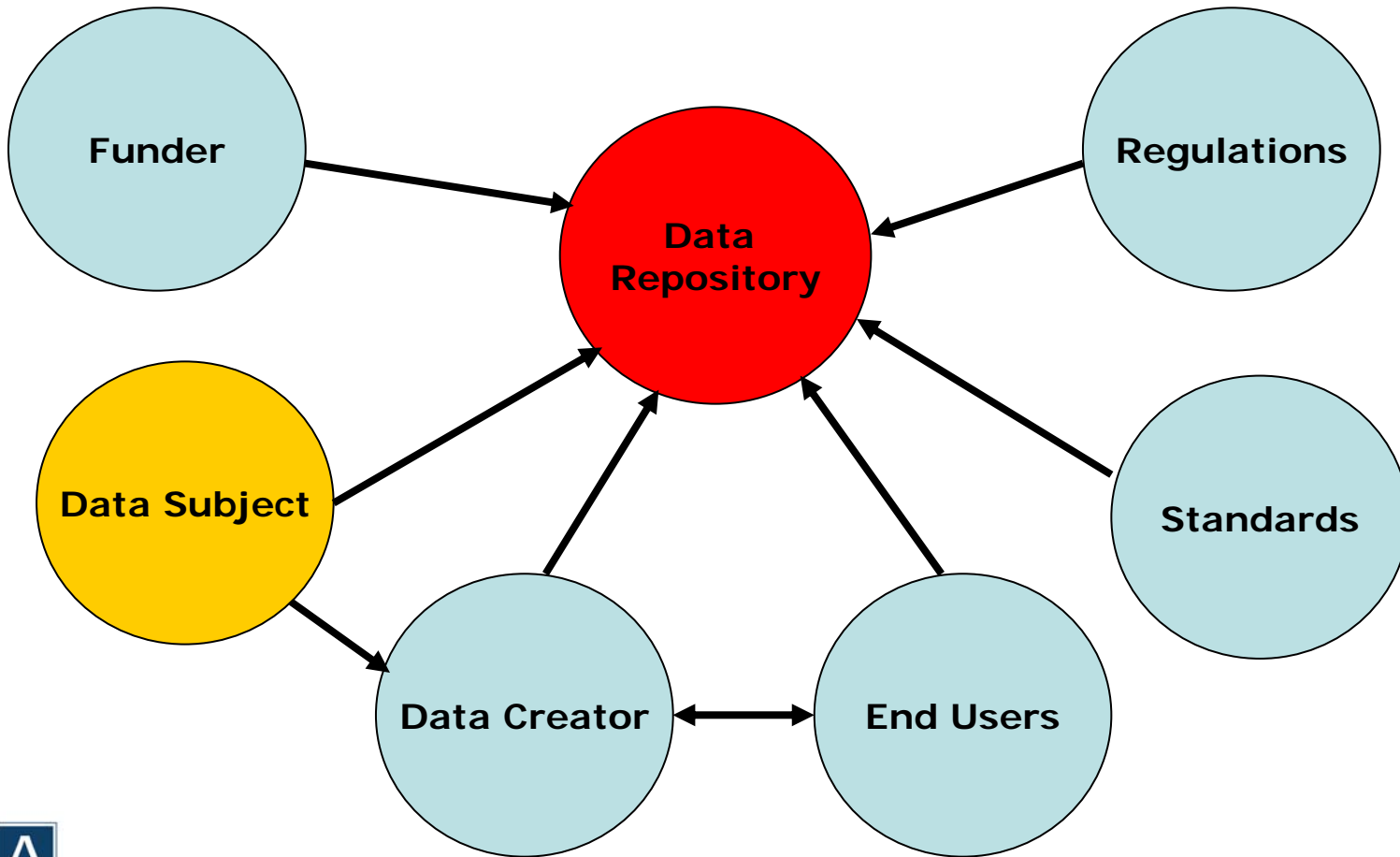
Reducing (preservation) costs

- Acting on evidence
- Including “basic” preservation type actions in Data Management Plans
- Examining new strategies for different access horizons (near term; long term)
- “Technical” solutions

Third-party agreements

- Responsibilities
 - Could possibly change (e.g., public records, change in user community)
- Clear and proportional assessment criteria
- Trust
 - All stakeholders including data subjects
 - Certification / Accreditation
 - Data Seal of Approval

Network of trust



Other issues

- Free-riding
- Who benefits from use? And how?
- Impact – measurement
 - Benefits / value
- Option strategy
- Rights initiatives may be stifling
- Hybrid funding models