

# RESEARCH DATA SHARING IN THE UK

Innovation and collaboration between funding  
councils, government, and academia

Funding provided by:



Service delivered by:



# UK UNIQUE IN EUROPE FOR ITS RESEARCH DATA PROVISION

- Data sharing policies
- Data curation infrastructure
- Cross-council collaboration
- Collaboration beyond academia (other research funders, other data creators)
- Partnership with National Statistical Institute
- Depth/breadth of data available
- Flexibility of data access regimes

# LARGER CONTEXT OF DATA SHARING IN SOCIAL SCIENCES : A TUG OF WAR

OPEN DATA

VS

PRIVACY

(with a sprinkling of budget cuts)

# DATA FEEDING FRENZY

Data Liberation Front

Cloud Computing

Mine it  
Merge it  
Mash it

Let's Google Map It!

**Publicly funded should mean publicly available**

## DATA PARANOIA

**Millions of benefits records lost**

**Laptop with official secrets  
left on train**

**Data CDs sent in post LOST**

**Identity Theft on the Rise**

**Census "resisters" organise**

# IMPORTANCE OF DATA SHARING

- Data collection is expensive and often unrepeatable
- Research councils/governments need to maximise value and impact of investments
- Cumulative knowledge building through subsequent use – science builds on itself
- Data sharing improves research quality -- keeps researchers / data creators honest, improves record-keeping and resource management

# ESRC AND DATA SHARING

- Data sharing policy – oldest in UK and with the biggest teeth
- Data curation and data sharing infrastructure – longest standing and most efficient in Europe
- Training in research data management and responsible data use
- Supporting researchers in secondary use of data
- Founder of UK Data Forum to ensure data resources and funding maximised, not duplicated

# UK DATA ARCHIVE

- 43 years experience in the field
- Initially established for holding ESRC-funded data and being conduit for supply of data internationally (ICPSR and CESSDA archives) – on magnetic reels!
- From 1970s distribution and preservation of data from government sources
- Runs a number of data-related services including ESDS, SDS, SQB, CRS and Census Portal, RELU, HDS, RDMS – some co-funded
- Is an official Place of Deposit for the UK National Archives
- Is a department of the University of Essex

# UK DATA ARCHIVE

- Thousands of active users, tens of thousands of datasets disseminated every year
- Beyond data itself: training and user support in specific substantive areas and methods
- Adding value to data – documentation, user guides, metadata records, teaching datasets, online data browsing tools, international data access
- Expertise in digital preservation, data management and data security (ISO 27001 and government accredited for IL2/IL3 data)

# COLLABORATION IS KEY

- Working with researchers creating data from the time they apply for funding to ensure data is shareable
- Working with government departments to ease the burden of data dissemination and minimise the risks of disclosure
- Collaborating with networks of data archives and services internationally to maximise access
- Working closely with the funding councils as expert advisor and collaborator on data sharing and data service issues
- Developing the research data service profession and impacting the data sharing landscape internationally (standards, best practice, technical developments, policies)

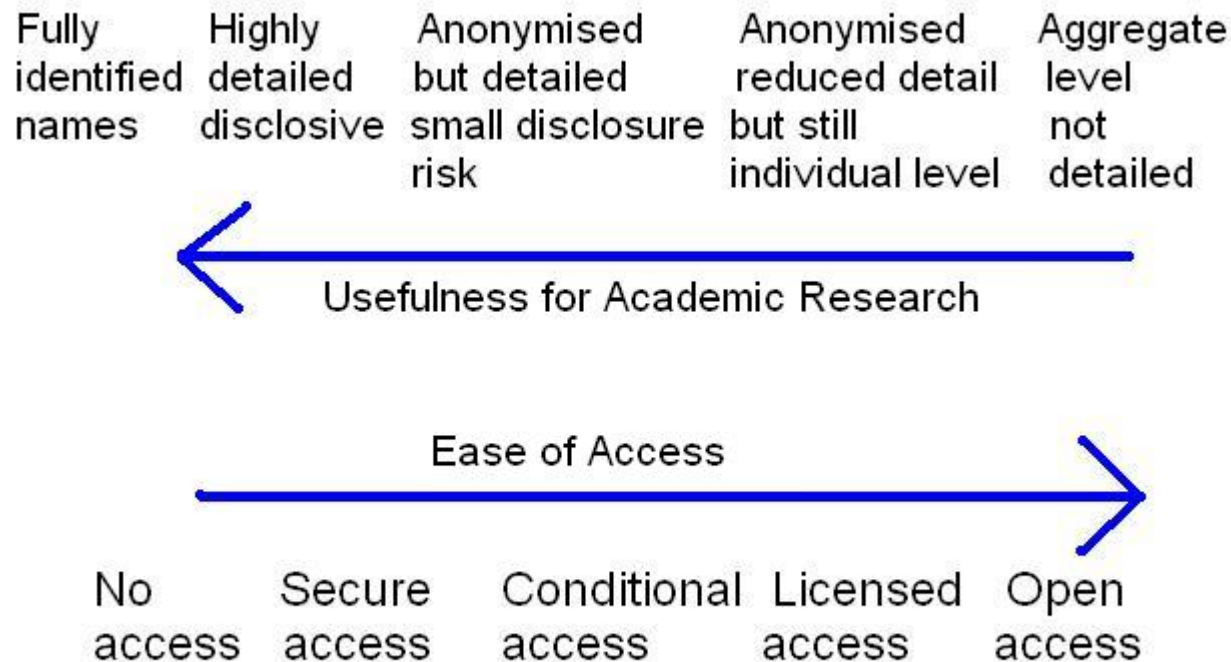
# DATA ACCESS FOR THE SOCIAL SCIENCES

Different access regimes for different kinds of data

- Open access / anonymous download (very few datasets) – Public Use data
- End User License (EUL) (vast majority of our 6000+ dataset collection) – Scientific Use data
- Special Conditions (small but significant proportion) – Scientific Use data, sensitive
- Special License / Approved Researcher (a few dozen datasets) – Restricted data
- Secure Data Service (a dozen and growing) – potentially disclosive data, previously only available in on-site enclaves

# THE DATA CONTINUUM

It's a trade-off between ease of access and usefulness for research



# MANAGING RISK



Data use is a risky business



So what we need are some good safety nets

# IT'S ALL ABOUT TRUST

- The most important security measure is to get the researchers to buy into data security as their own project
- Convenience is a positive security measure : eliminating the real-world causes of breaches (evidence-based security philosophy)
- Training is absolutely central : both how to do it right, why to do it right, and what the penalties are for doing it wrong
- Backed up by appropriate legal licensing framework and agreements
- Backed up by technology to first prevent and second identify misuse and provide reliable audit trails
- Backed up by commensurate penalties

# IT'S ALL ABOUT TRUST

- Collaborating with data creators early means a relationship of trust and mutual respect is developed.
- Flexible access regimes means data owners can trust the level of protection is appropriate
- Data value increases with use – we are the means to maximising the return on data investment
- Being a broker between data users and data creators, we facilitate trust and collaboration, not suspicion and control
- Win-win-win

# FUTURE OF SOCIAL SCIENCE DATA PROVISION

- ESRC bringing together all of its data-related services under one National Service umbrella in autumn 2012
- European social science data infrastructure being developed with EU funding (CESSDA ERIC and Data Without Boundaries project)
- More cross-council collaboration, e.g. new Birth Cohorts Facility
- Continued development of standards, best practice and new technologies

# CONTACT

Melanie Wright

Director, Secure Data Service

Associate Director, UK Data Archive

<http://securedata.data-archive.ac.uk>

<http://data-archive.ac.uk>

<http://esds.ac.uk>

melanie@essex.ac.uk