



DATA LIFE CYCLE & DATA MANAGEMENT PLANNING

VEERLE VAN DEN EYNDEN
RESEARCH DATA MANAGEMENT TEAM
UK DATA ARCHIVE
UNIVERSITY OF ESSEX

LOOKING AFTER AND MANAGING YOUR RESEARCH DATA (GOING DIGITAL AND ESRC ATN EVENTS)
UK DATA ARCHIVE, COLCHESTER, 24-25 APRIL 2013





BENEFITS OF MANAGING AND SHARING YOUR DATA

DATA CREATED FROM RESEARCH ARE VALUABLE RESOURCES THAT CAN BE USED AND RE-USED FOR FUTURE SCIENTIFIC AND EDUCATIONAL PURPOSES. SHARING DATA FACILITATES NEW SCIENTIFIC INQUIRY, AVOIDS DUPLICATE DATA COLLECTION AND PROVIDES RICH REAL-LIFE RESOURCES FOR EDUCATION AND TRAINING



THE LIFE OF DATA

A dataset has a longer lifespan than the research project that creates it

Data can be used and re-used for future research, if:

- shared
- managed well
- properly preserved
- made available



GOOD MANAGEMENT OF DATA

- good research
- high quality data
- data can be understood and used now and in future
- data can be shared and re-used
- needs to be planned
- needs to be specific for purpose



BENEFITS OF DATA SHARING

Public

- production of high quality research with social value

Funders

- make optimal use of publicly funded research
- avoid duplication of data collection

Scholarly community

- maximise transparency where appropriate
- quality improvement from verification, replication and trustworthiness
- valuable resources for teaching, e.g. methodology
- promote innovation – unintended, new uses of data

Research participants

- allow maximum use of their contributed data / information
 - minimise data collection on the hard-to-reach (e.g. ill, elites)
-

RESEARCH FUNDER DATA POLICIES

Research Councils UK Common Principles on Data Policy (May 2011)

- *Publicly funded research data are a public good, produced in the public interest, that should be made openly available with as few restrictions as possible in a timely and responsible manner that does not harm intellectual property.*
- in accordance with relevant standards and community best practice
- metadata to make research data discoverable
- legal, ethical, commercial constraints on release of research data
- recognition for collecting & analysing data; limited privileged use
- acknowledge sources of data, intellectual contributions, terms & conditions
- use public funds to support the management and sharing of publicly-funded research data



DATA LIFECYCLE & DATA MANAGEMENT PLANNING

A DATA MANAGEMENT AND SHARING PLAN HELPS RESEARCHERS CONSIDER: WHEN RESEARCH IS BEING DESIGNED AND PLANNED, HOW DATA WILL BE MANAGED DURING THE RESEARCH PROCESS AND SHARED AFTERWARDS WITH THE WIDER RESEARCH COMMUNITY

AREAS OF COVERAGE

- Data management planning tasks and the research lifecycle
- Data management checklist
- Roles and responsibilities
- Costing data management



WHY DATA MANAGEMENT PLANNING

- Research funders require planning for data management and data sharing, e.g. UK Research Councils
 - which data created during research
 - which policies apply (legal, institutional,....)
 - which data standards
 - how document data
 - ownership, copyright, IPR
 - data storage, backup, security
 - how share, preserve, incl. access
 - roles & responsibilities
 - costing resources needed

DCC: [UK research funders' DMSP expectations](#)



WHY DATA MANAGEMENT PLANNING

Research benefits

- think what to do with research data, how collect, how look after
- keep track of research data (e.g. staff leaving)
- identify support, resources, services needed
- plan storage, short & long-term
- plan security, ethical aspects
- be prepared for data requests (Fol, funder)

Overview of UK research funders with data sharing policies and their requirements for data management and sharing plans»

Funder	Plan required?	Required at application	Data topics in DMP
Arts and Humanities Research Council (AHRC)	Yes	Technical plan	Standards, preservation, continued access and use
Biotechnology and Biological Sciences Research Council (BBSRC)	Yes	Data management and sharing plan	Type, format, standards, sharing methods, restrictions, sharing timeframe
Cancer Research UK (CRUK)	Yes	Data sharing plan	Volume, format, standards, metadata, documentation, sharing method, timescale, preservation, restrictions
Department for International Development (DFID)	Yes	Access and data management plan	Repositories, limits, timescale, responsibilities, resources, access strategy
Engineering and Physical Sciences Research Council (EPSRC)	No	Policy framework at institutional level (from 2015)	
Economic and Social Research Council (ESRC)	Yes	Data management plan	Volume, type, quality, archiving plans, difficulties sharing, consent sharing, IPR, responsibilities
Medical Research Council (MRC)	Yes	Data management plan	Collection methods, documentation, standards, preservation, <u>curation</u> , security, confidentiality, sharing and access, timescale, responsibilities
Natural Environment Research Council (NERC)	Yes	Outline data management plan	Data management procedures, created data
Science and Technology Facilities Council (STFC)	Yes	Data management plan	Type, preservation, metadata, value, sharing, timescale, resources needed
<u>Wellcome Trust</u>	Yes	Data management and sharing plan	What data, when share, where share, how access, limits, how preserve, what resources



HOW

- Funder template for DMP
 - [ESRC DMP requirements in data policy](#) and [DMP guidance](#)
 - [MRC DMP guidance and template](#)
 - [AHRC technical plan requirements](#)
- DCC's [DMPonline](#) tool
- [UK Data Archive data management checklist](#)



PLAN YOUR DATA MANAGEMENT

- plan data management early
- assign roles and responsibilities
- design data management according to needs and purpose of research
- implement and review data management throughout research



ROLES & RESPONSIBILITIES

Assign, not presume roles or responsibilities for data management

Who?

- project director: design and oversee research
- research staff: design research; collect, process, analyse data; decide where data will be held, who will have access
- laboratory or technical staff: generate metadata / documentation
- database designer
- external contractors: data collection, data entry, transcribing, processing or analysis
- support staff: manage / administer research and research funding; ethical review; assess IPR
- institutional IT services: data storage, security, back-up services
- external data centres or archives: facilitate data sharing

COSTING

- Identify resources needed to make research data shareable beyond primary research team - above planned standard research procedures and practices
- Resources = people, equipment, infrastructure, tools to manage, document, organise, store and provide access to data
- Early planning can reduce costs

See our [data management costing tool](#)

ADDITIONAL ACTIVITY	COMMENTS AND SUGGESTIONS	✓	COST
<p>CONSENT FOR DATA SHARING Do you need to ask participants for their consent for data to be shared?</p> <p>Essential for qualitative interviews; possibly less so in quantitative surveys where data can be more easily anonymised.</p>	<ul style="list-style-type: none"> • when consent for data sharing is considered as part of standard consent procedures - very low or no additional cost • when participants need to be re-contacted or re-visited after research to obtain retro-active consent for data sharing - could be high cost • does this require preparation of information sheets, consent forms or extra time for consent discussions/staff training? 		
<p>DIGITISATION Do analogue or paper-based research data need to be digitised to increase their potential for sharing?</p>	<ul style="list-style-type: none"> • is additional equipment or software needed for conversion? • if simply image scanning text - may be low cost • if Optical Character Recognition required then checking for accuracy needed - may be high cost 		

DATA MANAGEMENT CHECKLIST

- points to relevant to consider when planning appropriate data management for research
- select what is relevant for your research
 - ✓ Are you using standardised and consistent procedures to collect, process, check, validate and verify data?
 - ✓ Are your structured data self-explanatory in terms of variable names, codes and abbreviations used?
 - ✓ Which descriptions and contextual documentation can explain what your data mean, how they were collected and the methods used to create them?
 - ✓ How will you label and organise data, records and files?
 - ✓ Will you apply consistency in how data are catalogued, transcribed and organised, e.g. standard templates or input forms?
 - ✓ Which data formats will you use? Do formats and software enable sharing and long-term validity of data, such as non-proprietary software and software based on open standards?
 - ✓ When converting data across formats, do you check that no data or internal metadata have been lost or changed?
 - ✓ Are your digital and non-digital data, and any copies, held in a safe and secure location?
 - ✓ Do you need to securely store personal or sensitive data?
 - ✓ If data are collected with mobile devices, how will you transfer and store the data?

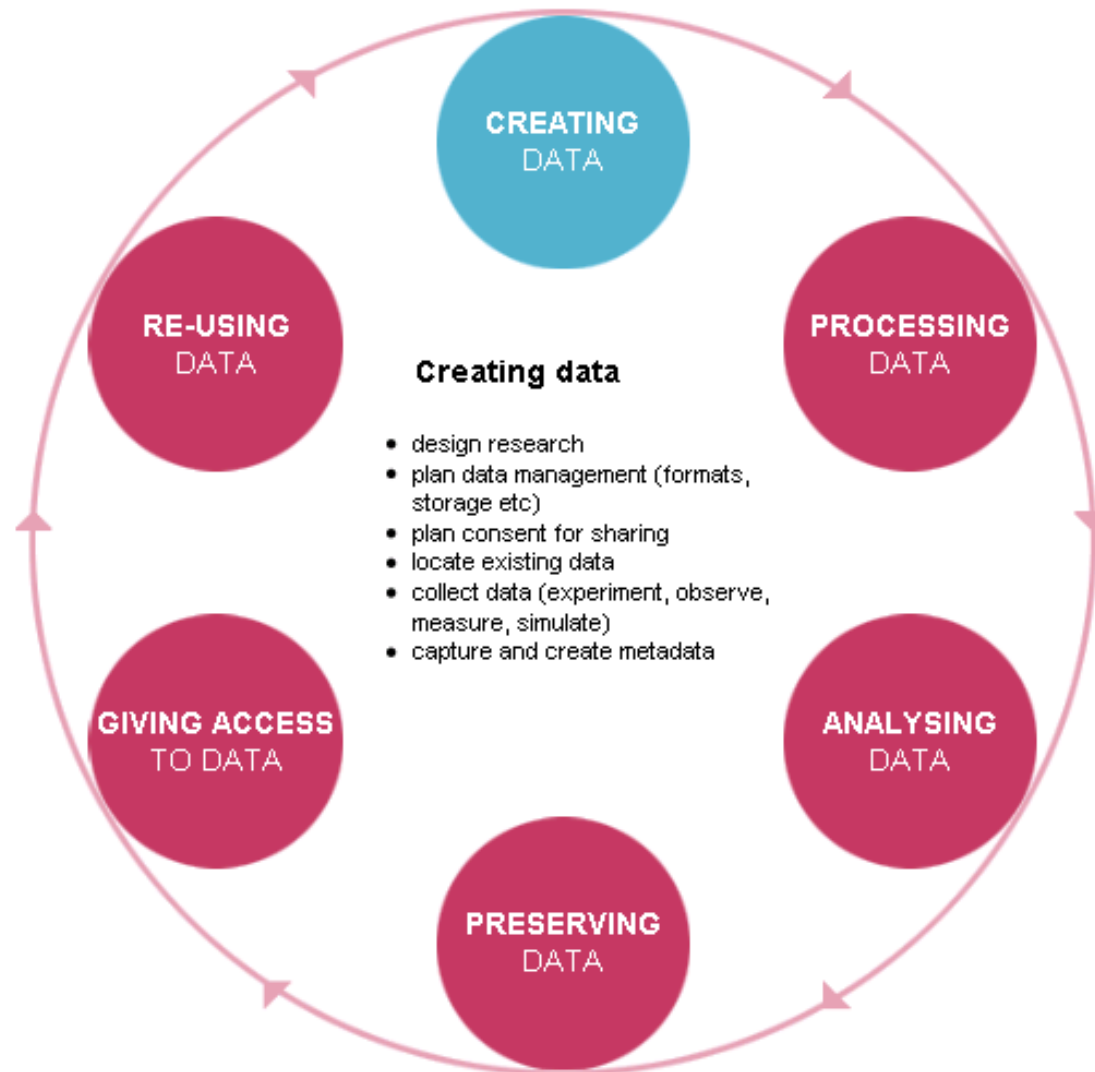
www.data-archive.ac.uk/create-manage/planning-for-sharing/data-management-checklist

PLANNING FOR RESEARCH ORGANISATIONS

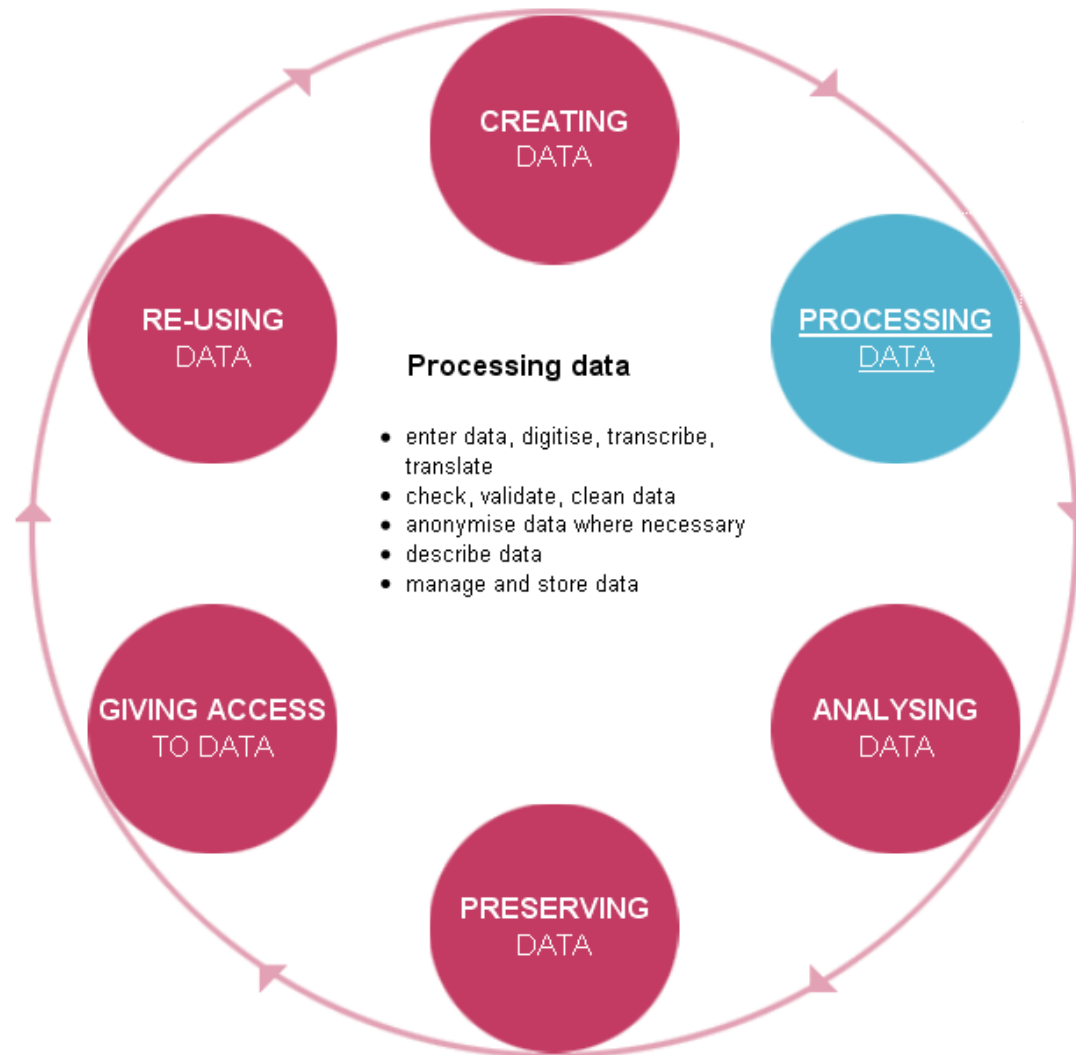
Organisations can provide framework of guidelines, tools and best practices to facilitate data management planning

- data management resource library
- data inventory for all projects
- assign data management responsibilities
- standardised forms, e.g. consent, ethical review
- transcription standards, incl. confidentiality agreement for transcribers
- file sharing and storage procedures
- security policy for storage, transmission of data
- data retention / destruction policy
- data copyright /ownership statement
- standard data format recommendations
- version control and file naming guidelines

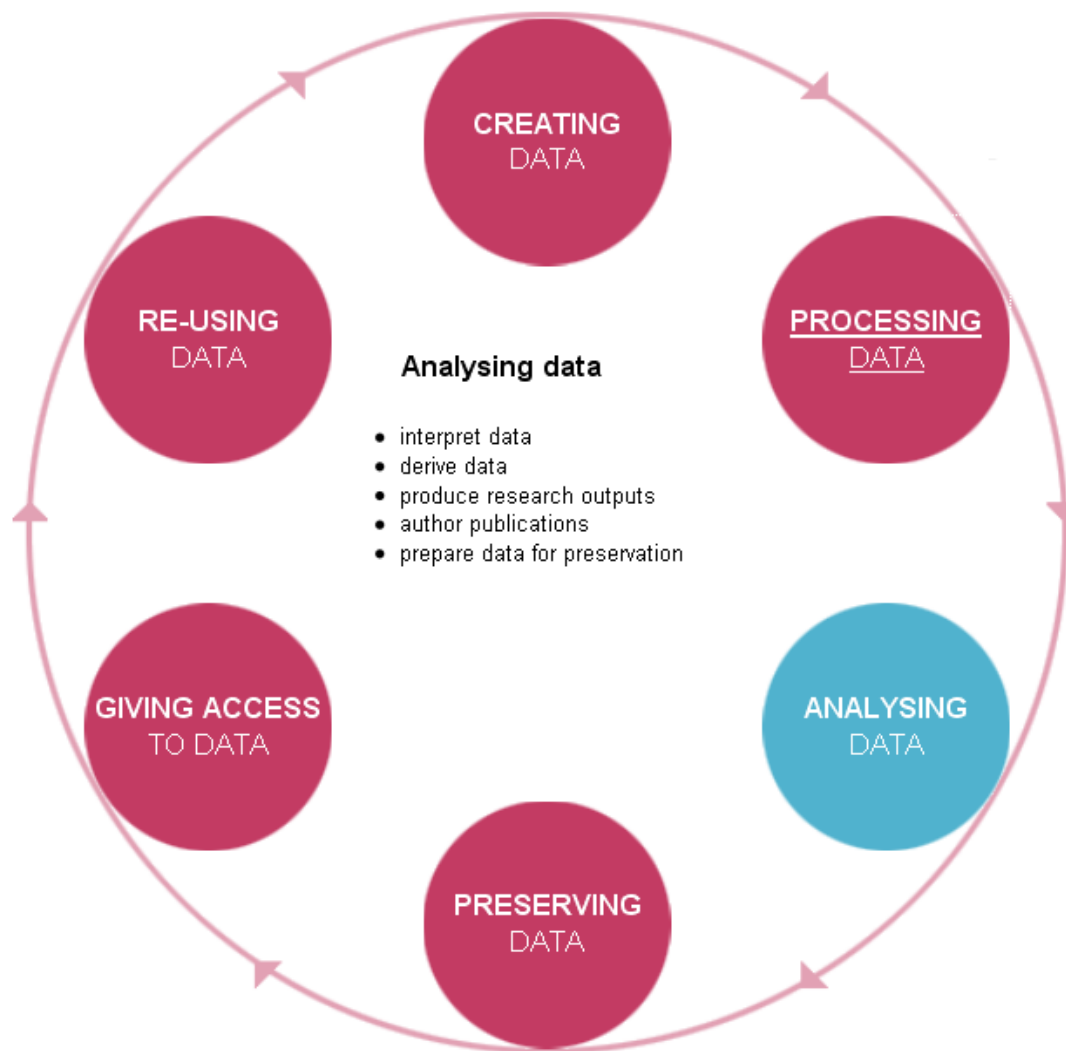
DATA LIFE CYCLE



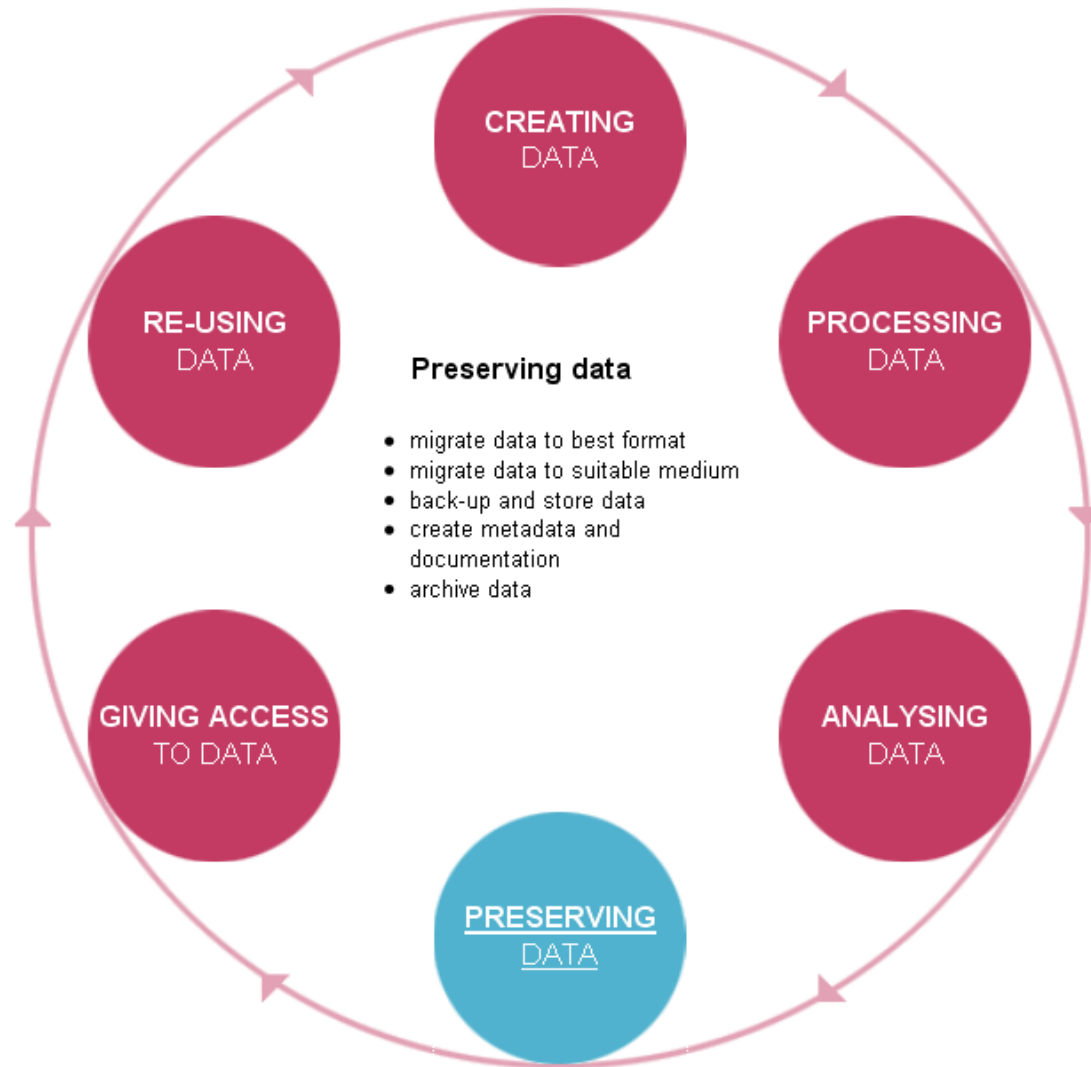
DATA LIFE CYCLE



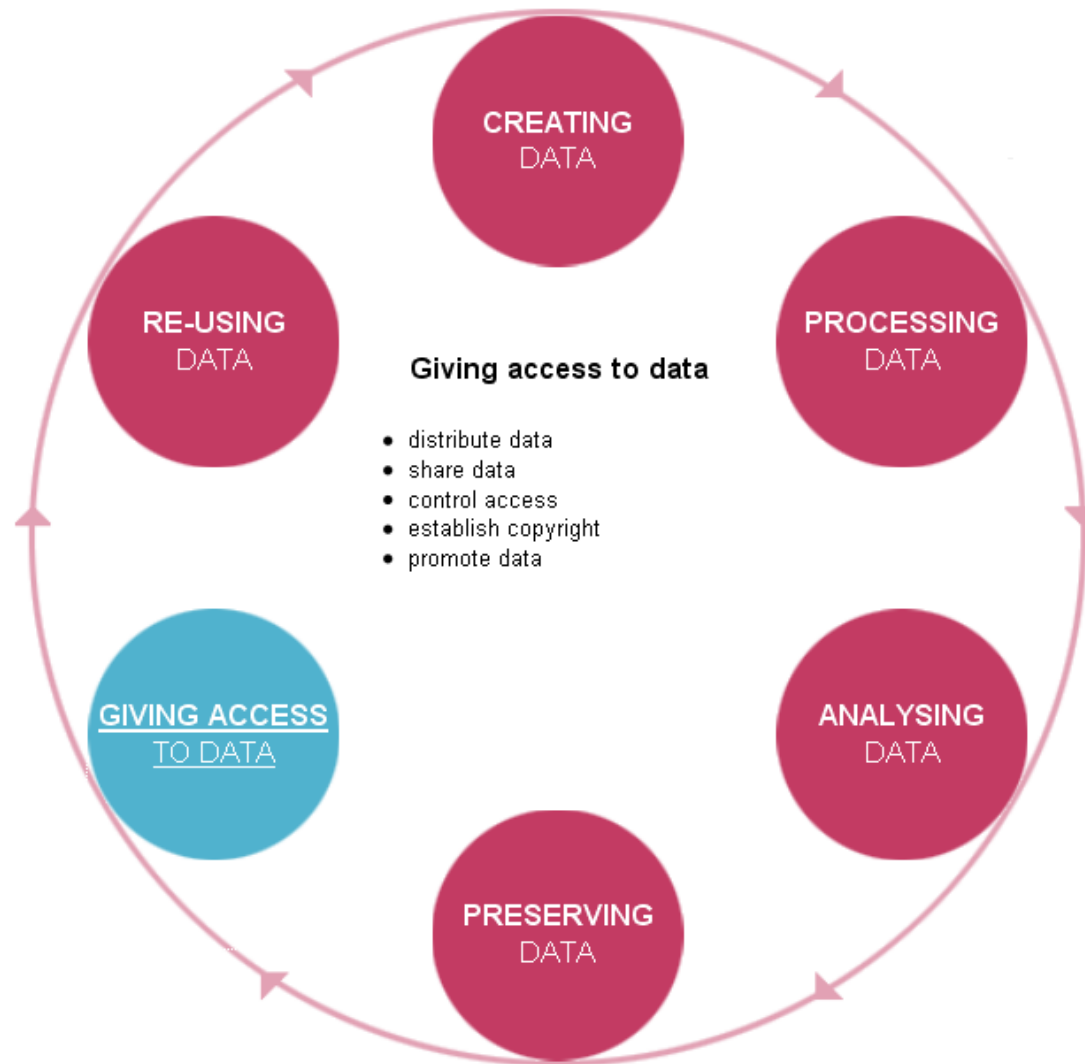
DATA LIFE CYCLE



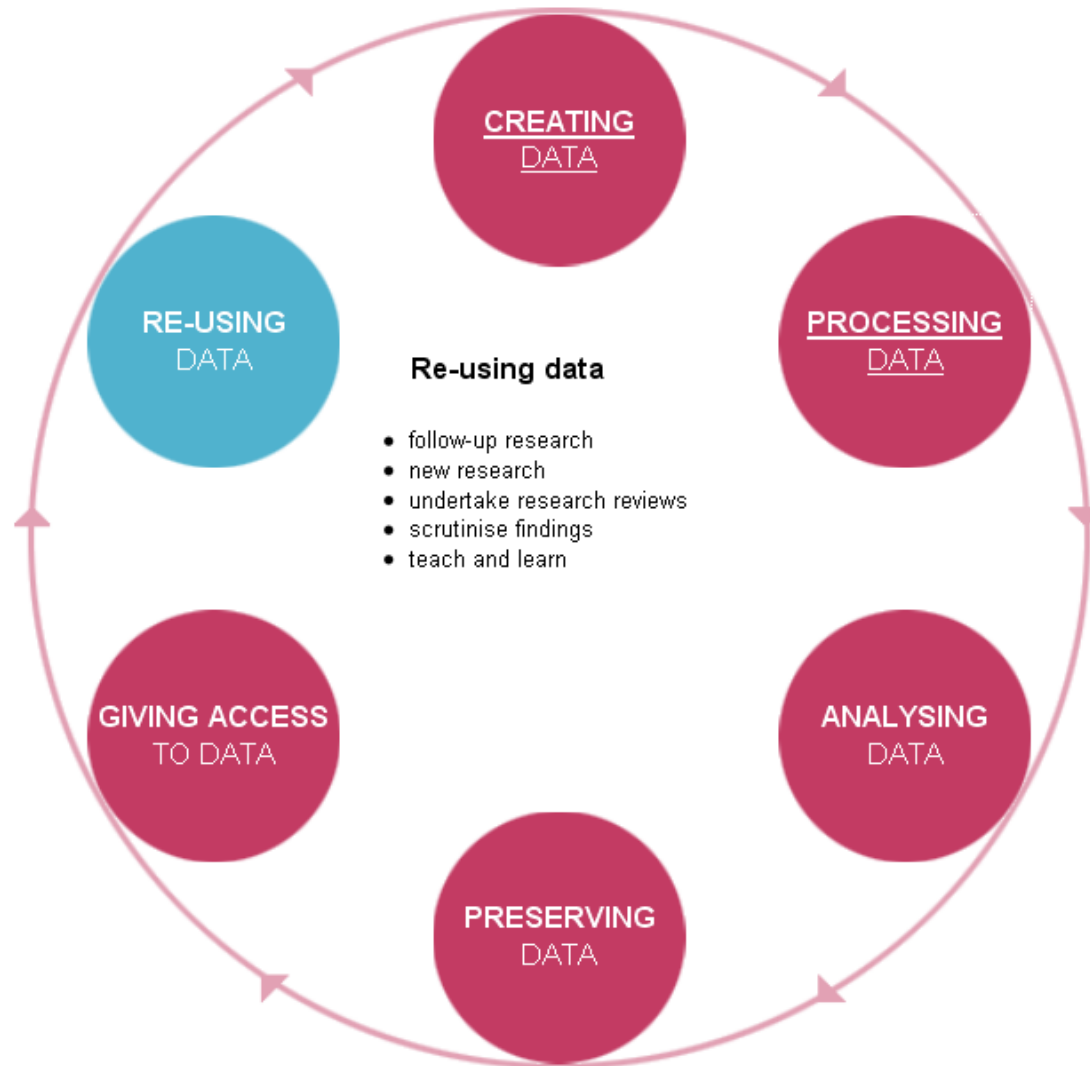
DATA LIFE CYCLE



DATA LIFE CYCLE



DATA LIFE CYCLE



EXAMPLE: HEALTH AND SOCIAL CONSEQUENCES OF THE FOOT AND MOUTH DISEASE EPIDEMIC IN NORTH CUMBRIA, 2001-2003 (SN5407)

Data re-used in study:
'Assessment of
Knowledge Sources
in Animal Disease
Control'

Research design
Consent for participation
& primary data use
Participants keep diaries
Interviews recorded

Interviews transcribed
Diaries transcribed
(MS Word)

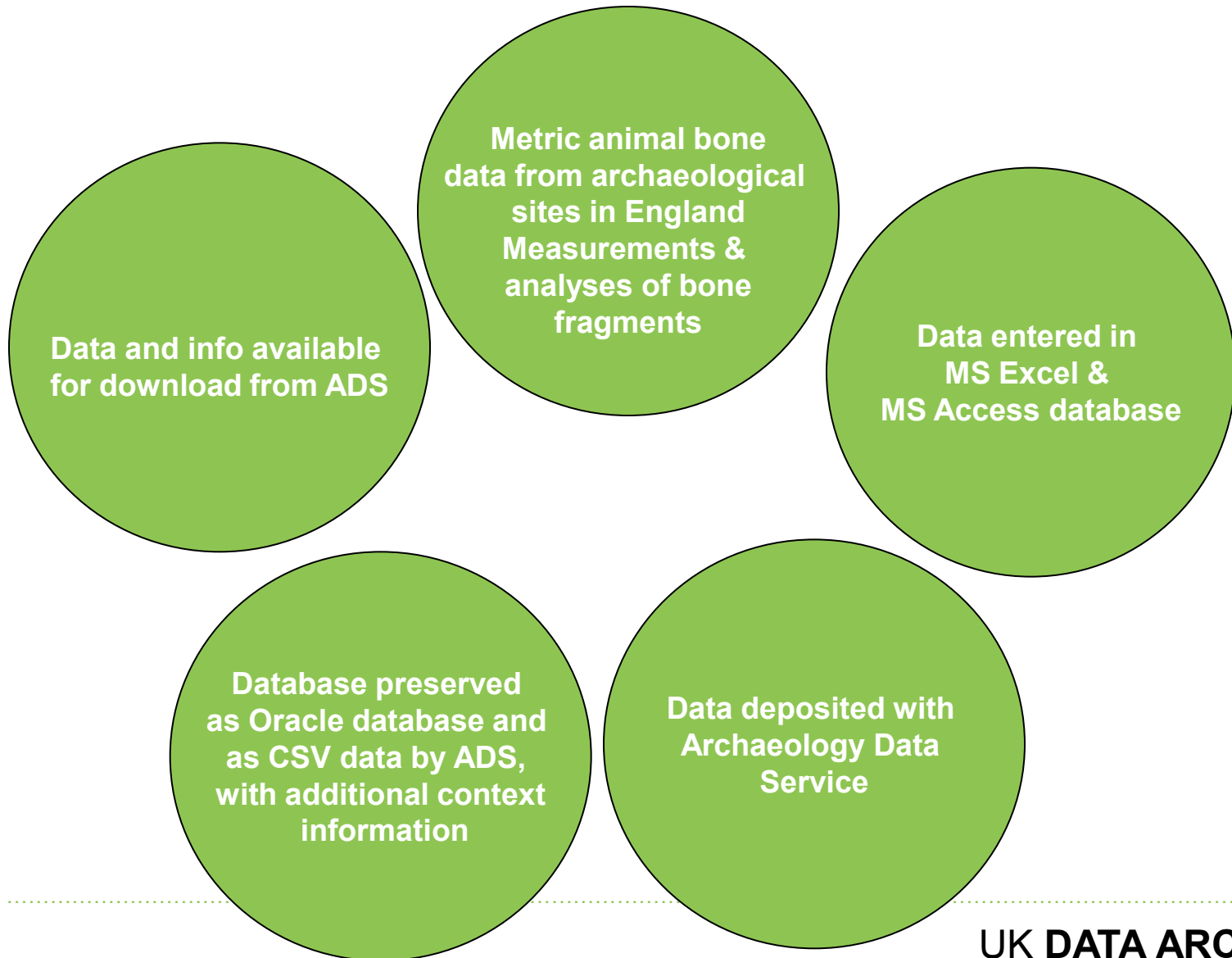
Transcripts
and user guide
available from UKDA

Transcripts and
recordings archived
at UKDA (RTF, MP3)
Catalogue record
created
User guide
created

Data archiving
discussed with
participants.
Consent to archive
transcripts
and recordings
obtained

UK DATA ARCHIVE

EXAMPLE: WESSEX ARCHAEOLOGY METRIC PROJECT





UK DATA ARCHIVE

- the UK Data Archive has over forty years experience in selecting, ingesting, curating and providing access to social science data
- we have huge experience of supporting researchers and data creators of social science data and related disciplines
- we do data sharing for the ESRC Data Policy (since 1995) and the Rural Economy and Land Use programme (2004-2012)
- our best practice approaches to making data shareable are based on:
 - challenges faced by researchers to share data
 - handling research data – quantitative and qualitative
- highly skilled staff comprising researchers, technical and information specialists

www.data-archive.ac.uk

OUR MANAGING AND SHARING DATA RESOURCES

Managing and sharing best practice guidance

- our website
- published guide
- Sage handbook (forthcoming Nov 2013)
- training programme

www.data-archive.ac.uk/create-manage

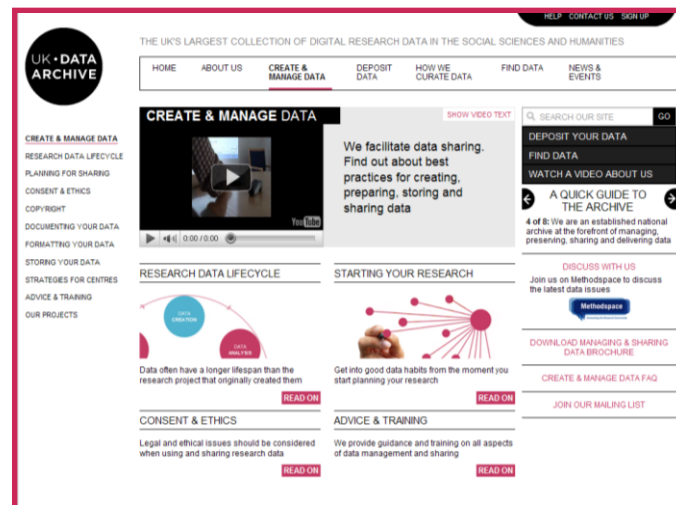
www.data-archive.ac.uk/media/2894/managingsharing.pdf



Training resources:

- presentations
- exercises and discussions / answers

www.data-archive.ac.uk/create-manage/training-resources



Planning for sharing



Consent and ethics



Copyright



Documenting your data



Formatting your data



Storing your data



Strategies for centres

SCENARIO

Producing urban asylum

- exploring experiences of the UK asylum dispersal system in different cities
- considering the impact of shifting policy regimes on asylum experiences
- methods:
 - archival research looking at policy documents and media coverage
 - ethnographic work carrying out observations of asylum activist groups and campaigns
 - 120 interviews with a range of asylum sector actors and asylum seekers themselves
- data:
 - interview recordings and interview transcripts (external transcriber)
 - observations – written notebook and transcribed
 - policy documents – transcribed
 - newspaper coverage – transcribed
 - online discussion thread ‘stories of asylum’

Prepare a data management plan

Where intervene in data lifecycle ?



CONTACT

UK DATA ARCHIVE
UNIVERSITY OF ESSEX
WIVENHOE PARK
COLCHESTER
ESSEX CO4 3SQ

T +44 (0)1206 872143
E: datasharing@data-archive.ac.uk
W: www.data-archive.ac.uk
