

Benefits of data sharing and basic data management

Northumbria University

Cristina Magder

Data Collections Development Manager

13 June 2022



Event Schedule

- 10:00 - 10:15 Introduction to UK Data Service and secondary data sources
- 10:15 - 10:30 Data Management Planning and FAIR data sharing
- 10:30 - 10:50 Ethical and Legal Considerations
- 10.50 - 11:00 Break
- 11:00 - 11.30 Data curation
- 11:30 - 11:40 Licence frameworks and reproducibility
- 11:40 - 12:00 Close and Q&A

Questions Padlet

Please see our Padlet for questions at

<https://padlet.com/dcmagd/y72xxi1tovwvvguu>

Introduction to UK Data Service and secondary data sources

UK Data Service (UKDS)

- UK's largest collection social, economic and population data
- funded by the Economic and Social Research Council (ESRC)
- provide users with access, support, guidance and training to facilitate high quality social and economic research and education
- support the development of best practices for data preservation and sharing standards participants

ukdataservice.ac.uk

UK Data Archive

- the lead partner of the UK Data Service
- centre of excellence in acquiring, curating and providing access to the largest collection of social science and population data for over 50 years
- provides a Trusted Digital Repository for the data you want stored

data-archive.ac.uk

Stats about UKDS

- **@8,500** datasets in the collection
- **@250** new datasets and new editions added each year
- **@45,000** registered users
- **@130,000** downloads worldwide p.a

UKDS Users

- Academic researchers and students
- Government analysts
- Charities and foundations
- Business consultants & data analysts
- Independent research centers & think tanks



[Source](#)

UKDS Data Sources

- National statistical authorities
- UK government departments
- Intergovernmental organisations
- Research institutes
- Individual researchers



UKDS Data Portfolio

UK Surveys

Large-scale government funded surveys

Longitudinal

Major UK surveys following individuals over time

International

Multi-nation aggregate databanks and survey data

Census

Census data 1971 – 2011

Business

Microdata and administrative data

Qualitative

Range of multimedia qualitative data sources

Data Catalogue

Data catalogue

Search our data catalogue... × Search

[Search tips](#) > [Variable and question bank](#) > [QualiBank](#) >

STUDIES (8671) SERIES (78)

Date from: Date to:

Refine date

Topic:

Data Type:

Access:

Country:

Reset filters

Displaying 1 - 10 of 8671 results

Page 1 of 868

Results per page:

10

Sort by:

Most recently released

SN 8922 [English Housing Survey, 2019: Housing Stock Data: Special Licence Access](#)

Ministry of Housing, Communities and Local Government

SN 6697 [Business Structure Database, 1997-2021: Secure Access](#)

Office for National Statistics

SN 8666 [Smart Energy Research Lab Observatory Data, 2019-2021: Secure Access](#)

Elam, Simon | Webborn, Ellen | McKenna, Eoghan | Oreszczyń, Tadj | Anderson, Ben | Few, Jessica | Pullinger, Martin | European Centre for Medium-Range Weather Forecasts | Ministry of Housing, Communities and Local Government | Royal Mail Group Limited

SN 855574 [Impact of COVID-19 on Recent Graduate Career Decisions and Outcomes, 2020-2021](#)

Tomlinson, M, University of Southampton

UKDS Data Catalogue

Browsing Data

Browse by theme

Ageing >	COVID-19 >	Crime >	Economics >
Environment and energy >	Education >	Ethnicity >	Food >
Health >	Housing >	Information and Communication >	Labour >
Politics >	Poverty >		

[UKDS: Browsing Data](#)

Browse by data type

UK Surveys >	Cross-national surveys >	Longitudinal studies >	International macrodata >
Census data >	Business microdata >	Qualitative/mixed methods >	Administrative data >

UK Data Service Repositories

- The UK Data Service curated repository is most suitable for large-scale, population representative data. Curated collections include the Annual Population Survey, the Crime Survey for England and Wales, the English Health Survey, the Understanding Society and the Centre for Longitudinal Data cohorts as well as others; you can read more about our key data owners and data collections [here](#)
- The UK Data Service ReShare self-deposit repository is suitable for smaller-scale data; it was initially funded to host ESRC funded data (ESRC mandates data sharing) but currently hosts data funded by a various numbers of research councils as well as non-funded data within remit, and suitable to our designated user community

Licencing Framework

Open

- No real disclosure risk. Under open licences; almost no restrictions on reuse

Safeguarded

- Zero to low real disclosure risk. Requires authentication and authorisation e.g. registered user and End User Licence Agreement

Controlled

- Real disclosure risk. Requires project approval, user vetting and training; access via a safe setting; output checking

Three Tier Access Framework

Open Access

- Registration not required

Safeguarded Access

- Registration required
- Special conditions might apply

Controlled Access

- Registration required
- Access available via the [Five Safes Framework](#)

Registration

- Students and members of staff at a UK institution of higher or further education (UK HE/FE) or members of organisations part of the [UK Access Management Federation](#) (UKAMF), can [register with us](#) using the username and password issued by their institution
- Any other researchers can register with us by [applying for a username](#)

[Registration and login FAQs](#)

Why Is Registration Required?

- Registration and agreeing to the access conditions are part of the **ethical responsibility** in secondary data analysis
- The conditions of use require researchers to **act responsibly and ethically** with the data

[UK Data Service End User Licence Agreement](#)

Five Safes Framework



- **Safe data:** treated and de-identified
- **Safe projects:** approved by data owners/Research Accreditation Panel and must demonstrate public good
- **Safe people:** complete the Safe Researcher Training and gain Accredited Researcher status
- **Safe settings:** environment prevents unauthorised access i.e. UKDS SecureLab
- **Safe outputs:** screened and approved to ensure no disclosive information leaves the Safe Settings

Benefits of Depositing with UK Data Service

- no costs associated with depositing
- an online platform for creating, uploading, and managing deposits
- a three-tier licensing framework facilitating a robust access policy
- data management guidance and support provided to ensure legal sharing and re-use
- resource discovery and citation supported with persistent identifiers (DOIs)
- new deposits promoted to national and international researchers via newsletters, partner archives and research organisations

Support from UK Data Service

- [Events](#)
- [YouTube Channel](#)
- [Learning Hub](#)
 - [Research Data Management](#)
 - [New to Using Data](#)
 - [Data Skills Modules](#)
 - [Teach with Real Data](#)
- [ReShare Review Procedure](#)
- [Get in touch](#)

Data Management Planning and FAIR data sharing

Data Lifespan and Data Sharing

- data have a longer lifespan than the (research) project that creates them
- data can be (re) used for various purposes when well managed, preserved and shared
- who is benefiting from data sharing?
researchers, public, funding bodies, scholarly community and research participants

What are the benefits of data sharing?

Take a few minutes to consider this and add it to the Padlet.

<https://padlet.com/dcmagd/r66o60xqj0g5saz8>

Benefits of Sharing Data

Researchers	Funding Bodies	Public	Research Participants
increases visibility of scholarly work and increase citations (DOI)	make optimal use of publicly funded research	advance science to the benefit of society	allow maximum use of their contributed data/information
archiving provides long-term safe storage for data	avoid duplication of data collection	adoption of emerging norms – ‘open access’ publishing	minimise data collection on the hard-to-reach (e.g. ill, elites)
assists in implementing publishers’ data retention policies and enables collaborations	maximise return for investment	compliance with laws and regulations (transparency, openness, accountability)	enable participants’ experiences to be understood as widely as ethically possible

FAIR Principles for Publishing Data

Findable

Accessible

Interoperable

Re-usable

Force 11 FAIR Principles

How do you make your data FAIR?

Take a few minutes to consider this and add it to the Padlet.

<https://padlet.com/dcmagd/uv8peyuvinpau1rl>

Making Data FAIR

Findable

- community-endorsed metadata standards
- persistent identifiers e.g. DOIs
- responsible repositories

Accessible

- availability statement (including restrictions)
- methods/tools to access the data
- metadata preserved indefinitely

Interoperable

- standard vocabularies/ontologies
- standard metadata schemas

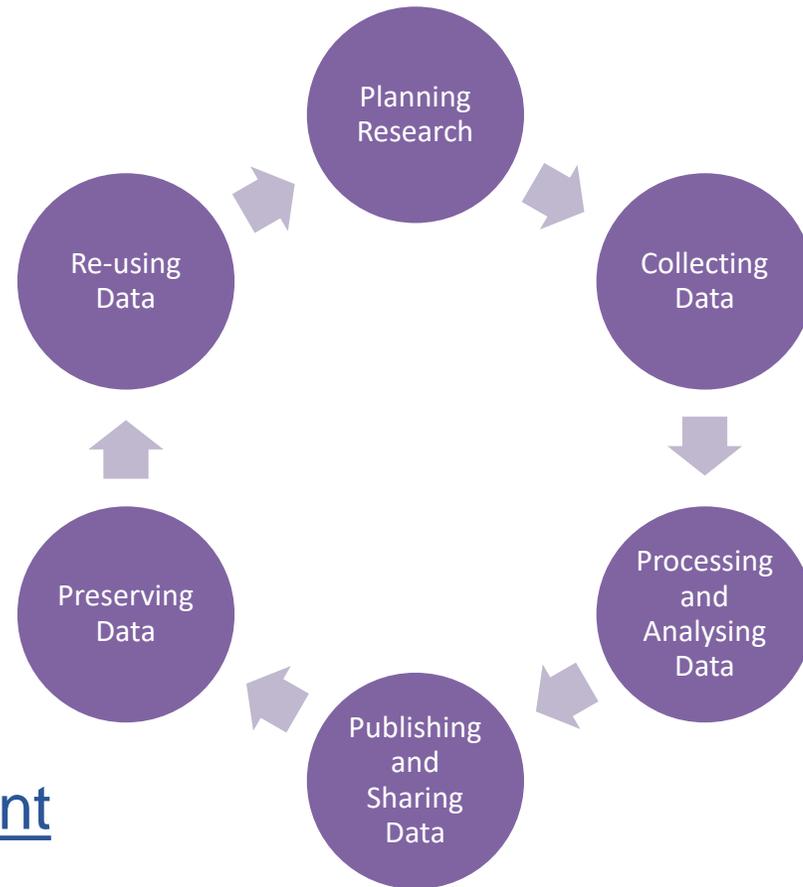
Re-useable

- community-endorsed data licencing
- established data quality assurance processes
- long-term preservation

Data Management Planning

A data management plan (DMP) describes how the data will be:

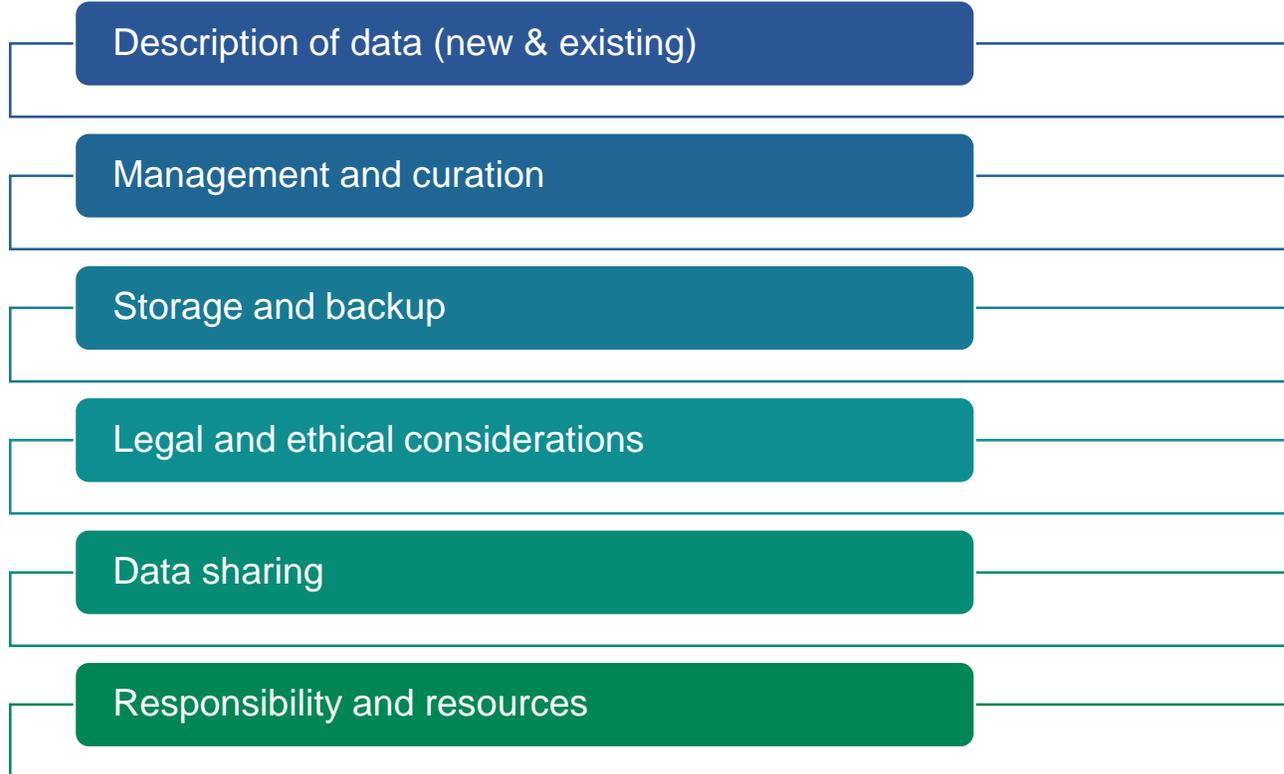
- Collected
- Organised
- Analysed
- Preserved
- Shared



[UK Data Service Data Management Plan Checklist](#)

DMPs Structure

Data management plans can differ from one funder/institution to another - always consult your funder/host institution requirements



DMP Structure from UK Research Funders

Funder	DMP required?	Required at application	DMP topics
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, timescales, responsibilities, resources, access strategy
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, preservation, standards, curation, security, confidentiality, sharing and access, timescale
Natural environment research council (NERC)	Yes	Outline data management plan	Data management procedure, created data
SCEINCE & Technology Facilities Council (STFC)	Yes	Data management plan	Type, preservation, metadata, value, sharing, timescale, resources needed
Wellcome Trust	Yes	Data management and sharing plan	What data, when share, where share, how access, limits, how preserve, what resources

Data management checklist (UKDS)

Ethical and Legal Considerations

with many thanks to Hina Zahid

Key Principles for Ethical Research (ESRC & UKRI)

- To maximise benefit for individuals and society & minimise risk and harm
- The rights and dignity of individuals and groups should be respected
- Voluntary and appropriately informed participation
- Research should be conducted with integrity and transparency
- Clearly defined lines of responsibility and accountability
- Independence of research should be maintained and where conflicts of interest cannot be avoided they should be made explicit

Ethical Considerations in Data Sharing



- Clear guidance designed by the National Statistician's Data Ethics Advisory Committee's (NSDEC)
 - Ethical Principles
 - Ethics Self-Assessment Tool

Duty of Confidentiality and Data Sharing

- Exists in UK common law and may apply to research data
- Disclosure of confidential information is lawful when:
 - the individual to whom the information relates has consented
 - disclosure is necessary to safeguard the individual, or others, or is in the public interest
 - there is a legal duty to do so, for example a court order
- Best practice is to avoid very specific promises in consent forms

Data Protection Legislation Considerations

If personal data about people is collected or used in research data protection regulations applies

- Data Protection Act 2018 (DPA), General Data Protection Regulation 2018 (GDPR) & the UK GDPR 2021
- DPA (2018) & the UK GDPR (2021) applies when
 - ✓ a researcher based in the UK collects personal data about people anywhere in the world
 - ✓ a researcher outside the UK collects personal data on UK citizens
- DPA (2018), EU GDPR (2018) & the UK GDPR (2021) applies when
 - ✓ a researcher based in the UK collects personal data about people across Europe

Personal Data

Personal data is any information relating to an identified or identifiable natural person. This information can directly identify individuals such as names, home addresses, IP addresses or indirectly identify individuals by combining the information that is made available

Special considerations are given to sensitive personal data referred to as special category data such as ethnicity, race, political or religious beliefs, genetic, biometric and health data, trade union membership, sex life and sexual orientation information

Principles of Processing Personal Data

(DPA and the UK GDPR)

All data must be:

- be processed lawfully, fairly and transparently
- be kept to the original purpose
- be minimised (only the personal data that is necessary is collected)
- be removed if not necessary
- have the accuracy upheld
- be kept confidential and integrity maintained

Consent, Anonymization and Access Control

- Make explicit mention of the planned procedures to handle consent for data sharing for data obtained from human participants, and/or how to anonymise data, and how the data will be accessed to make sure that data can be made available and accessible for future scientific research
- If you are unsure of how issues of confidentiality are to be addressed to facilitate data sharing, please [get in touch](#) for advice

Consent in Research

- Consent for research ethics: provide information regarding study purpose, risks, benefits, voluntary participation
- Consent can also be used as a legal basis for the processing of personal data under GDPR

How to Seek Consent

- Consent can be gained in written or oral form
- Format depends on the kind of research
- Important to document how it has been gained, what information has been provided to the participants and what they have agreed to

Consent Documentation: Information Sheet

An information sheet should cover the following topics:

- Purpose of the research
- What is involved in participating
- Benefits and risks of participating
- Procedures for withdrawal
- Usage of the data during research, dissemination, storage, publishing and archiving
- Details of the research: funding source, sponsoring institution, name of project, contact details for researchers, how to file a complaint

Consent Documentation: Consent Form

Consent form should:

- Use simple language and free from jargon
- Allow the participant to clearly respond to points such as:
 - The participant has read and understood information about the project
 - The participant has been given the opportunity to ask questions
 - The participant voluntarily agrees to participate in the project
 - The participant understands that they can withdraw at any time without giving reasons and without penalty
 - Future uses (e.g. publications, share and reuse)
 - Signatures and dates of signing for the participant and the researcher

Further Resources for Consent

- [UKDS Model Consent Form](#)
- [Example Consent Forms](#)
- [Example Information Sheet](#)
- [Consent for data sharing](#)
- [DARIAH ELDAH Consent Form Wizard | CFW](#)

Copyright Considerations

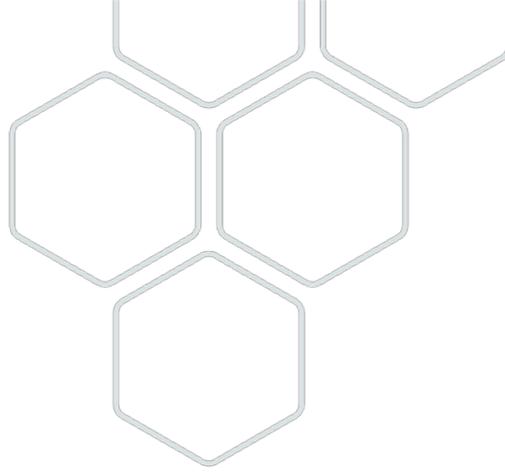
- Copyright is an intellectual property right assigned automatically to the creator
- Data owner (researcher) has copyright of research data
- Compiled datasets contain original copyright – seek permission to archive when collecting
- Data archives publish data – hold no copyright
- If information is in the public domain (e.g. online) it does not mean copyright does not apply!

Difficulties in sharing data

- Identify any potential obstacles to sharing your data and explain possible measures you can apply to overcome these
- State explicitly which data may be difficult to share and why
- If ethical issues could cause difficulties in data sharing, explain your strategies for dealing with these issues

Further information

- Guidance on ethical issues
- Guidance on rights in data
- Guidance on Data Protection Act
- Guidance on other rights including Freedom of Information Act



Thank you.

