UK Data Service

Data management basics: Ethical and legal issues in data sharing

Presenter: Dr Hina Zahid

Senior Research Data Officer

UKRI Economic and Social Research Council

Overview

- Section 1: Key Principles for Ethical Research.
- Section 2: Research Data and Legal Frameworks.
- Break.
- Section 3: Consent in Research.
- Section 4: Key Considerations in Secondary Data Use.
- Q&A Session.

Learning objectives

1. Be able to identify key principles for ethical research.

2. Understand the legal frameworks that govern research data.

3. Understand the role of consent in research and data sharing.

4. Understand key considerations when using secondary data.

Research data
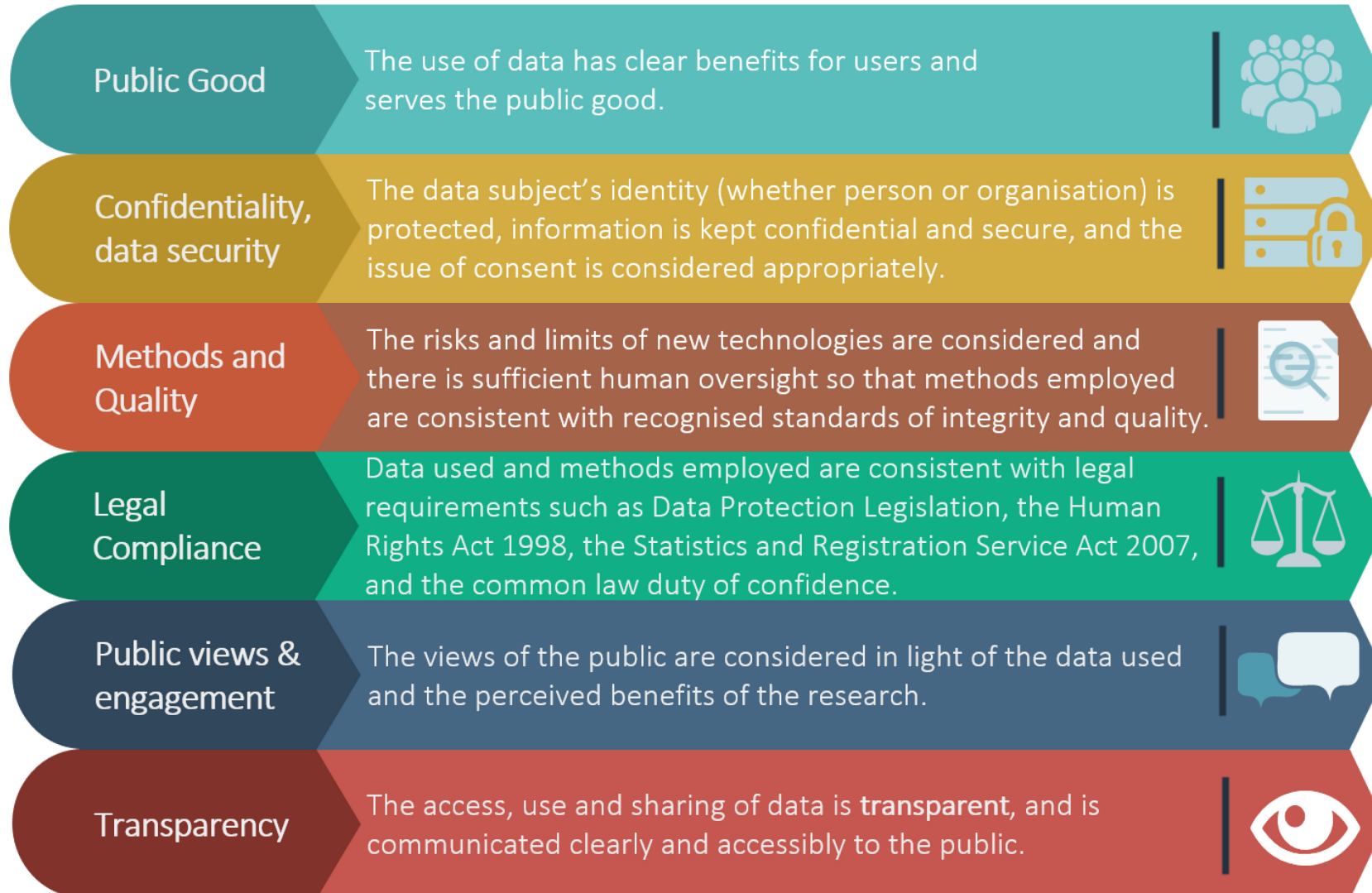
Primary data
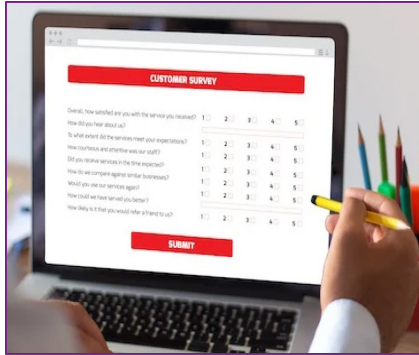
Secondary data

Key principles for ethical research

- To maximise benefits 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).

- UKSA [Ethics Self-Assessment Tool](#).

NSDEC ethics principles

Public Good
The use of data has clear benefits for users and serves the public good.

Confidentiality, data security
The data subject's identity (whether person or organisation) is protected, information is kept confidential and secure, and the issue of consent is considered appropriately.

Methods and Quality
The risks and limits of new technologies are considered and there is sufficient human oversight so that methods employed are consistent with recognised standards of integrity and quality.

Legal Compliance
Data used and methods employed are consistent with legal requirements such as Data Protection Legislation, the Human Rights Act 1998, the Statistics and Registration Service Act 2007, and the common law duty of confidence.

Public views & engagement
The views of the public are considered in light of the data used and the perceived benefits of the research.

Transparency
The access, use and sharing of data is **transparent**, and is communicated clearly and accessibly to the public.

Best practices for ethical sharing of research data

- Ethical obligations should be considered throughout the research lifecycle.

- Be knowledgeable about relevant research organisations own standards and requirements.

- Comply with relevant laws.

- Avoid social and personal harm.

- Data centres facilitate ethical and legal re-use of research data, protection of participants and safeguarding of personal data.

Research data and legal frameworks

Types of data
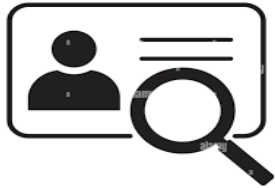
1. Data that relates to identified or identifiable individuals.

2. Anonymous data that no longer relates to identified or identifiable individuals.

3. Synthetic data.

Personal information

- Personal information or data is any information relating to an identified or identifiable natural person (name, address, postcode).

- It also includes sensitive personal data or special category data such as ethnicity, political or religious beliefs, biometric data, health.

- If personal information about people is collected or used in research data protection regulations apply.

Research Data 1: Data that relates to identified or identifiable individuals

- The common law duty of confidentiality.

- Data protection legislation (the UK General Data Protection Regulation (UK GDPR) within the Data Protection Act (DPA) 2018).

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 considerations

If personal information about people is collected or used in research data protection regulations apply.

EU GDPR (2018) & country specific laws:

- a researcher based in a country from the European Union (country that adheres to EU GDPR) collects personal data about people living in a country from the European Union.

e.g. DPA (2018) & the UK GDPR 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 .

e.g. DPA (2018), EU GDPR (2018) & the UK GDPR applies when

- a researcher based in the UK collects personal data about people across Europe.

Data subjects rights

- The right to be **informed.**
- The right of **access.**
- The right to **rectification** (correction).
- The right to **erasure** (right to be forgotten).
- The right to **restrict processing.**
- The right to data **portability.**
- The right to **object.**

Data protection legislation: DPA & UK GDPR

Principles of data processing

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.

Legal bases to process personal data

Legal base	Example
Consent	Survey to capture public opinion, whereby email addresses are collected to contact respondents at a later stage. Qualitative study on a sensitive topic, e.g. violence against women, where respondents may be identifiable from the collected information. Oral history project where people's real names are used.
Public interest / public task	Longitudinal study of people living with dementia and their carers, to identify how people would like to be supported. Findings inform and support the caring strategy and public advocacy
Legitimate interest	Research project funded and undertaken by a private corporation to look at the effects of smoking on car passengers.
Protect vital interests	Unlikely in research. Hospital treating a patient after a serious road accident can search for his/her ID to find previous medical history or to contact his next of kin.
Legal obligation	Unlikely in research. Processing personal data as part of a health and safety report or incident.
Performance of a contract	Unlikely in research. Processing personal data as part of an employment contract.

Legal basis to process personal data

- Research conducted in the UK:

    - **Task in public interest** (for all public bodies (NHS / HSC, Universities, UKRI, etc).

    - **Legitimate interest** (for all non-public bodies (charities, commercial companies, etc).

- If you are processing special category data::

    - Legal basis.

    - Additional condition .

    - Data protection impact assessment (DPIA) for any type of processing which is likely to be high risk.

GDPR and research

- Principles.
- Rights of data subjects.
- Processing grounds for processing personal data.
- Emphasis on transparency, clear information, clear documentation.
- Reuse for research allowed with safeguards.

Research Data 2: Anonymous data that no longer relates to identified or identifiable individuals

Effective anonymization	Pseudomisation
Application of one or more anonymisation techniques to personal information. When done effectively, the anonymised information cannot be used by the recipient to identify the data subject either directly or indirectly	Processing of personal data in such a manner that it can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures
Not subject to the UK GDPR	Subject to the UK GDPR

Further information:
- UKDS guidance on How to anonymise quantitative and qualitative data
- ICO guidance on anonymisation, pseudonymisation

Research Data 3: Synthetic Data

- Synthetic data is information that is artificially created (algorithmically) rather than generated by real-world events.

- Not subject to data protection legislation or the common law duty of confidentiality.

Consent in research

Consent is the process by which a researcher discloses appropriate information about the research so that a participant may make a voluntary, informed choice to accept or refuse to cooperate.

Why to seek consent?

- Consent for research ethics.
- Consent for processing of personal data.

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.

Three key areas to be addressed

Wording in consent forms and information sheets could be broken down in **three** key areas:

- Taking part in the study.

- Use of the information in the study.

- Future use and reuse of the information by other.

Challenges in obtaining informed consent

- Participant perception and expectations.

- Children & vulnerable people.

- Participant's poor awareness of their rights.

- Failure to provide adequate information.

- Absence of consideration of participant's background such education, culture.

- Use of jargon.

- Sceptical of confidentiality issues.

Handling personal data in research

- Will the research project collect personal data or special categories of data?
- Who will be the data controller for the research project?
- Will the research involve collaboration with other partners.
- Who will have access to the personal data collected?
- What ground will be used for processing the personal data in a project?
- What information needs to be communicated to participants?
- How and where will the personal data be stored?

Strategies for managing and sharing research data obtained from people

- Protection of identities when promised (anonymisation, de-identification).

- Processing ground for personal data (consent).

- Regulated access where needed (open, safe guarded, controlled).

Managing access to data

Open
- Available for download/online access under open licence without any registration.

Safeguarded
- Available for download / online access to logged-in users who have registered and agreed to an End User Licence (e.g. not identify any potentially identifiable individuals).
- Special agreements (depositor permission; approved researcher).
- Embargo for fixed time period.

Controlled
- Available for remote or safe room access to authorised and authenticated users whose research proposal has been and who have received training.

Key Considerations in Secondary Data Use

What rights might there be in research data?

Copyright

Database rights

Further information on rights in research data and different scenarios

Rights in research data

- Right granted to creators and owners of works that are the result of human intellectual creativity.

- Assigned automatically to the creator.

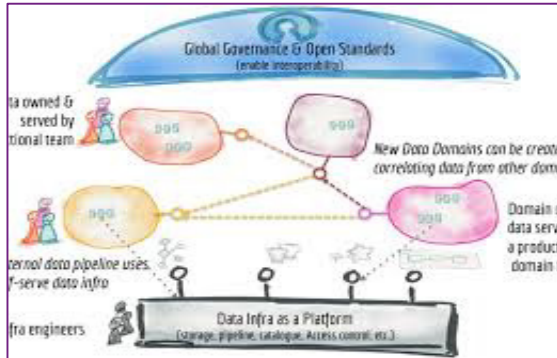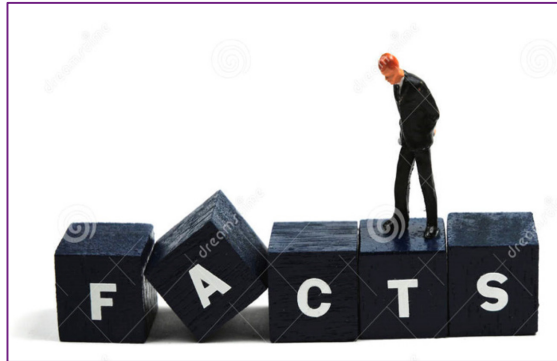- When data is shared, original owner is the right holder.

Further information on rights in research data and different scenarios

When my research data be protected by copyright?

When data is created, shared & reused

When my research data be protected by copyright? Cont…

What actions require permission?

- Permission is not required:
  - Incorporation of the factual data in one's own words in a structure of one's own.
  - making a copy (including a digital copy) and utilizing that copy for one's own research, as long as it is not made available to others.
  - Citing from the research data.

- Permission is required:
  - Inclusion of the research data in a publication.
  - Sharing the research data with other people.
  - Including the whole of the research data in a database of one's own that is also shared with other people.

Copyright and personal data

Where personal data is concerned:

- Permission from the person who has created the work is required (the copyright holder).

- Permission from all the people whose personal data is in the work is required. This includes opinions about people.

Best practices when using secondary data

Questions to ask:
- Who the copyright holder of the datasets is?
- Are you allowed to use them and in what way?
- Are you allowed to archive and publish them in a data repository?

- If not, you may need to seek for further permission to distribute material you do not own - copyright clearance.

- If permission is not granted, need to remove copyrighted variables/material before publishing or sharing.

Resources

- [Copyright (UK Data Service)](#)
- [Variable information log template for secondary data users](#)

Best practice for legal compliance

- Investigate early which laws apply to your data, including cross-country collaborative working.
- Do not collect or keep personal or sensitive data if not essential to your research.
- Plan early on; seek advice from your research office.
- Ensure that you check participants know how this data will be used.
- Remember: not all research data are personal (e.g. anonymised data are not personal).

Further resources

- UK Data Service
- UKDS Model Consent Form
- Example Consent Forms
- Example Information Sheet
- Consent for data sharing
- DARIAH ELDAH Consent Form Wizard | CFW
- Rights when using secondary data sources
- Regulating access to data
- Managing and sharing research data: A guide to good practice

Thank you.

Dr Hina Zahid

hzahid@essex.ac.uk

datasharing@ukdataservice.ac.uk

Twitter: @hinazahid2016