



# ISSUES IN MANAGING VIDEO DATA

---

---

**LIBBY BISHOP**  
UK DATA ARCHIVE  
UNIVERSITY OF ESSEX

---

**WORKSHOP ON RESEARCH USING VIDEO**  
OPEN UNIVERSITY, 19 JULY 2013





# VIDEO THROUGH THE DATA LIFE CYCLE

---

## Topics for today's talk

- Brief history
- Funders' data policies
- Informed consent for video
- Anonymisation
- Controlling access to data
- Storage and transfer
- Dissemination

## BRIEF HISTORY OF VIDEO IN SOCIAL RESEARCH

---

- 1955 - First combined verbal and non-verbal – Center for Advanced Study of Behavioral Sciences – Stanford University
- 1967 – Video of middle-school children – Erickson
- 1970's –patient-doctor interactions – MSU – Jordan/Robillard
- 1980's – use of copier machines – Xerox PARC – Suchman
- 80s-90s – London Underground control room – Heath
- 1990's – classroom studies in maths – Lampert
- 1986 – 2000 - Institute for Research on Learning
- 2000's – multimodal studies – language/science – Goldman, Pea, Kress, Jewitt
- Erickson, F. Uses of video in social research: a brief history, IJSRM 14(3) 2011, 179-189.

# DATA PLAN REQUIREMENTS

---

Funder	Required at application	Data topics in DMP
AHRC	Technical plan	Standards, preservation, continued access & use
BBSRC	Data management and sharing plan	Type, format, standards, sharing methods, restrictions, timeframe
CRUK	Data sharing plan	Volume, format, standards, metadata, documentation, sharing method, timescale, preservation, restrictions
DFID	Access and data management plan	Repositories, limits, timescale, responsibilities, resources, access strategy
EPSRC	Policy framework	
ESRC	Data management plan	Volume, type, quality, archiving plans, difficulties sharing, consent sharing, IPR, responsibilities

# DATA PLAN REQUIREMENTS

---

Funder	Required at application	Data topics in DMP
MRC	Data management plan	Collection methods, documentation, standards, preservation, curation, security, confidentiality, sharing & access, timescale, responsibilities
NERC	Outline data management plan	DM procedures, created data
STFC	Data management plan	Type, preservation, metadata, value, sharing, timescale, resources needed
Wellcome Trust	Data management and sharing plan	What data? When share? Where share? How access? Limits, how preserve? What resources?

Digital Curation Centre, Funders' data plan requirements: [www.dcc.ac.uk/resources/data-management-plans/funders-requirements](http://www.dcc.ac.uk/resources/data-management-plans/funders-requirements)

Knight, G; (2012) Funder Requirements for Data Management and Sharing. London School of Hygiene and Tropical Medicine, London. [researchonline.lshtm.ac.uk/208596/](http://researchonline.lshtm.ac.uk/208596/)

---

## ETHICAL ARGUMENTS *FOR* ARCHIVING DATA

---

- Store and protect data securely
- Not burden over-researched, vulnerable groups
- Make best use of hard-to-obtain data (e.g., elites, socially excluded, over-researched)
- Extend voices of participants
- Provide greater research transparency
- Enable fullest ethical use of rich data

*In each, ethical duties to participants,  
peers and public may be present*



## DATA PROTECTION ACT, 1998

---

- Personal data:
  - relate to living individual
  - individual can be identified from those data or from those data and other information (i.e., any video showing faces)
  - include any expression of opinion about the individual
- Requirements for handling personal data
  - processed fairly and lawfully
  - obtained and processed for a specified purpose
  - adequate, relevant and not excessive for the purpose
  - accurate
  - not kept longer than necessary
  - processed in accordance with the rights of data subjects, e.g. right to be informed about how data will be used, stored, processed, transferred, destroyed, ...; right to access info and data held
  - kept secure
  - not transferred abroad without adequate protection
- Personal data can be disclosed only with consent

# DATA PROTECTION ACT & RESEARCH

---

- Exceptions for personal data collected as part of research:
  - can be retained indefinitely (if needed)
  - can be used for other purposes in some circumstances
  - people should still be informed
- **If data are anonymised (personal identifiers removed) then DP laws will not apply as these no longer constitute 'personal data'**

*DPA is not intended to, and does not, inhibit ethical research*

---





## OPTIONS FOR SHARING CONFIDENTIAL DATA

---

- Obtain **informed consent**, also for data sharing and preservation / curation
- Protect identities e.g. **anonymisation**, not collecting personal data
- **Regulate access** where needed (all or part of data) e.g. by group, use, time period
- Securely store personal or sensitive data

# CONSENT NEEDED ACROSS THE DATA LIFE CYCLE

---

- Engagement in the research **process**
  - decide who approves final versions of transcripts
- **Dissemination** in presentations, publications, the web
  - decide who approves research outputs
- Data **sharing** and archiving
  - consider future uses of data

Always dependent on the research context – special cases for covert research, verbal consent, etc.

# A GOOD INFORMATION SHEET & CONSENT FORM

---

- Meets requirements of Data Protection laws
  - purpose of the research
  - what is involved in participation
  - benefits and risks
  - mechanism of withdrawal
  - usage of data – for primary research and sharing
  - strategies to ensure confidentiality of data (anonymisation, access,....) where this is relevant
- Need to balance
  - as simple as possible
  - complete for all purposes: use, publishing, sharing
  - avoid excessive warnings
- UK Data Archive model consent form  
<http://www.data-archive.ac.uk/media/210661/ukdamodelconsent.doc>

## INFORMED CONSENT FOR UNKNOWN FUTURE USES

---

- It is possible to provide much information about reuse
  - who can access the data – only authenticated researchers
  - purposes – research or teaching or both
  - confidentiality protections, undertakings of future users
  - general consent (similar to consent with emergent research topics)
- Medical research and biobank models – enduring, broad, open consent
  - no time limits; no recontact required
  - unspecified hypotheses and procedures
  - 99% consent rate (2500+ patients) – Wales Cancer Bank

*ESRC expects that others will also use it [data], so consent should be obtained on this basis and the original researcher must take into account the long-term use and preservation of data. (ESRC Framework for Research Ethics, 1.17.5.1)*

## ANONYMISING QUALITATIVE DATA

---

- plan or apply editing at time of transcription  
*except: longitudinal studies - anonymise when data collection complete (linkages)*
- avoid blanking out; use pseudonyms or replacements
- avoid over-anonymising - removing/aggregating information in text make them unusable, unreliable or misleading
- consistency within research team and throughout project
- keep anonymisation log of all replacements, aggregations or removals made – keep separate from anonymised data files
- Video – usually best NOT to anonymise; for most image data, modifications sufficient to anonymise will damage integrity of data

## REGULATING ACCESS DATA

---

- UK Data Archive has gradation of access controls
  - small number of studies are open (no registration)
  - majority require registration
  - data users sign legally binding End User Licence – e.g. not identify any potentially identifiable individuals
  - stricter regulations for certain types of data:
    - Special Licences
    - Approved researchers
    - require data access authorisation from data owner prior to data release
    - embargo for given time period
    - Secure Data Service (no direct data access)
- Multiple AC can apply to different data types within one study

# DATA STORAGE

---

All digital media are fallible

File formats and physical storage media become obsolete

- optical (CD, DVD) and magnetic media (hard drive, tapes) degrade
- never assume the format will be around for ever

Best practice

- use data formats with long-term availability
- **storage strategy - at least two forms of storage and locations**
- maintain original copy, external local copy and external remote copy
- copy data files to new media two to five years after first created
- check data integrity of stored data files regularly (checksum)
- know your personal/institutional back-up strategy
- know data retention policies that apply: funder, publisher, home institution
- what to protect? not only data, and not only digital

# ENCRYPTION

---

Always encrypt personal or sensitive data

Encrypt anything you would not send on a postcard

- for moving files e.g. transcripts
- for storing files e.g. shared areas, mobile devices

Basic principles

- use an algorithm to transform information ( $A=1$ )
- need a 'key' to decrypt

Free softwares that are easy to use

- Safehouse
- Truecrypt
- Axcrypt

These softwares

- encrypt hard drives, partitions, files and folders
- encrypt portable storage devices such as USB flash drives





## DEMO: DATA ENCRYPTION

---

Create an encrypted storage space using free software SafeHouse

[www.data-archive.ac.uk/media/312652/storingyourdata\\_encryptionexercise.pdf](http://www.data-archive.ac.uk/media/312652/storingyourdata_encryptionexercise.pdf)

# FILE SHARING & COLLABORATIVE ENVIRONMENTS

Sharing data between researchers and teams

- too often email attachments
- Yousendit, Dropbox – consider if appropriate as services can be hosted outside the EU (DPA for personal data), e.g. encrypt
- Virtual Research Environments
  - MS SharePoint
  - Sakai
- file transfer protocol (ftp)
- physical media
- [Essex ZendTo](#)



By David Fletcher  
<http://www.cloudtweaks.com/2011/05/the-lighter-side-of-the-cloud-data-transfer/>

## OPTIONS FOR DEPOSITING DATA AT UKDA

---

- ESRC Data Store – repository for ESRC funded researchers
  - Current file size limited to 1G – plans to increase
  - We recommend alternatives for larger files
  - Can embargo for up to 18 mos
- Offers from other new depositors, no limit (in theory) but
  - Must meet Collections Development Policy
  - Current dissemination of video limited

## DISSEMINATING VIDEO -

---

- At UK Data Archive – large audio (and prospective video) by DVD as they are too large for download (we have to consider users with low bandwidth)
- <https://vimeo.com/> Free
- <http://www.smugmug.com> or <http://www.zenfolio.com/> gives more control, longer video (20 min) unlimited storage and branding for a £200 a year
- YouTube
  - + cheap and easy
  - - “lossy” format (lower quality) and no access controls
  - For data at UKDA – can link YouTube url to catalogue record

# CASE STUDY – VISUAL QUALITATIVE DATA 1

---

## Projects' backgrounds

- Connected Lives-Manchester/Leeds Real Life Methods
- Investigated how social networks and communities are perceived and represented in a heterogeneous inner city neighbourhood
- A Place Called Happiness – on dementia
- Walking interviews, participatory mapping, day-diaries
- Visuals: networks, researcher and participant photos
- Not designed to archive data – novel file formats, too complex

## Project addressed ethical issues with visual data

- Used iterative, flexible, situated ethical stance
- Gained verbal consent at onset of each interaction and during
- Requested permission to use outputs for dissemination
- Researchers regarded participants as “capable of engaging in a collaborative and negotiated research process”
  - Clark, A. (2012) in Pink, S. (ed.) *Advances in Visual Methodology*, Sage
- Participants also adjusted their methods of taking photos by
  - Avoiding taking images of people
  - Specifying varied uses for different materials



**UK DATA ARCHIVE**

## NEGOTIATING ETHICS

### “Show us around your neighbourhood”

#### The place of extraordinary events



F2: *[This] is where the riots actually started, well they actually started from t'middle o' road down here. So I don't know if you want to start from t'middle o'road. Are we going to go from where H~ got its notorious name, do you want to go from here?*

AC: *Ok*

F2: *So like from [19]95 from the riots we kind of got labelled this bad area and things changed.*



## CASE STUDY – VISUAL QUALITATIVE DATA 2

---

Some researcher decisions and participants' experiences

- Nothing used without consent from anyone identifiable
- Participants generally willing for materials to be used in displays
- Those who attended events were ok seeing themselves on public display
- In some cases, did not use images, even with consent (e.g., illegal acts)
- Found participants familiar with issues of what to display from Facebook, Flickr, and YouTube

Current reflections on feasibility of archiving this data

- Would participants have consented to archiving?
  - Probably yes for anonymised transcripts; not images
- Anonymisation – not preferred; damaging to visual data
- Context – visual materials seem to “need” more. Is this true?
- Archiving does not enable participants to negotiate forms of reuse. Does this matter?
- Broader concern – ethics compliance may systematically exclude certain research from archives: visual, covert, evaluation, critical. To what effect?

**UK DATA ARCHIVE**



## CASE STUDY – VISUAL QUALITATIVE DATA 3

---

*“I know it’s hard, but that doesn’t means we shouldn’t try.”*

Andrew Clark on archiving visual data

Based on papers, presentations and reflective notes by Andrew Clark (University of Salford) and gratefully used with his permission.



# CONTACT

---

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

---

T: +44 (0)1206 872001

E: [datasharing@ukdataservice.ac.uk](mailto:datasharing@ukdataservice.ac.uk)

W: [ukdataservice.ac.uk](http://ukdataservice.ac.uk)

---

**UK DATA ARCHIVE**