Who owns ‘your’ research data?: Legal and ethical issues in using and sharing data

Libby Bishop
Senior Officer Research Data Management Support Services

Bethany Morgan Brett
User training and support

University of Essex

22nd October 2013

UK Data Service

Who we are

- 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 creators of social science data
- we run workshops on data management issues, which are part of good research practice and also make data shareable

Data management topics

- sharing data
- data management planning
- documenting data
- formatting data
- data security and storage
- encryption, and file sharing
- ethics and consent
- data copyright

Our managing and sharing data resources

Training Resources:

- presentation PowerPoints
- exercises and answers

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

Managing and sharing guidance

- sections
- references
- training programme

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

UK Data Service

Benefits of managing and sharing your data

UK Data Service

Exercise: why is good data management important?

Write down five reasons why good data management is important?

1)
2)
3)
4)
5)

Benefits of good data management

- efficiency – makes research easier
- safety – protect valuable data
- quality – better research data = better research
- reputation – enhances research visibility
- compliance – with ethical codes, data protection laws, journal requirements, funder policies

Data lifecycle and data management planning

UK Data Service

The life of data

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

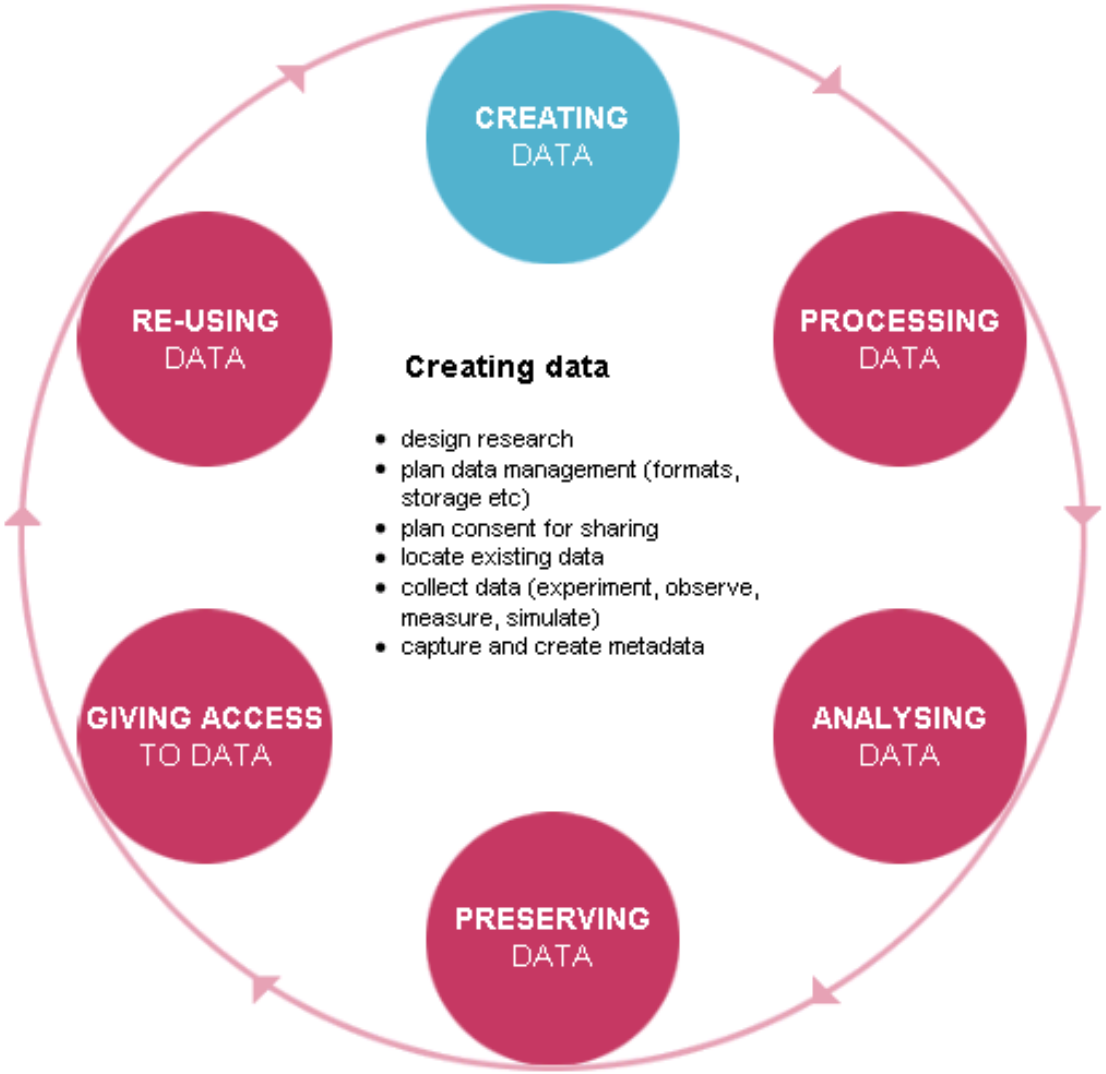
Data can be used and re-used for future research, if:

- shared
- managed well
- properly preserved
- made available

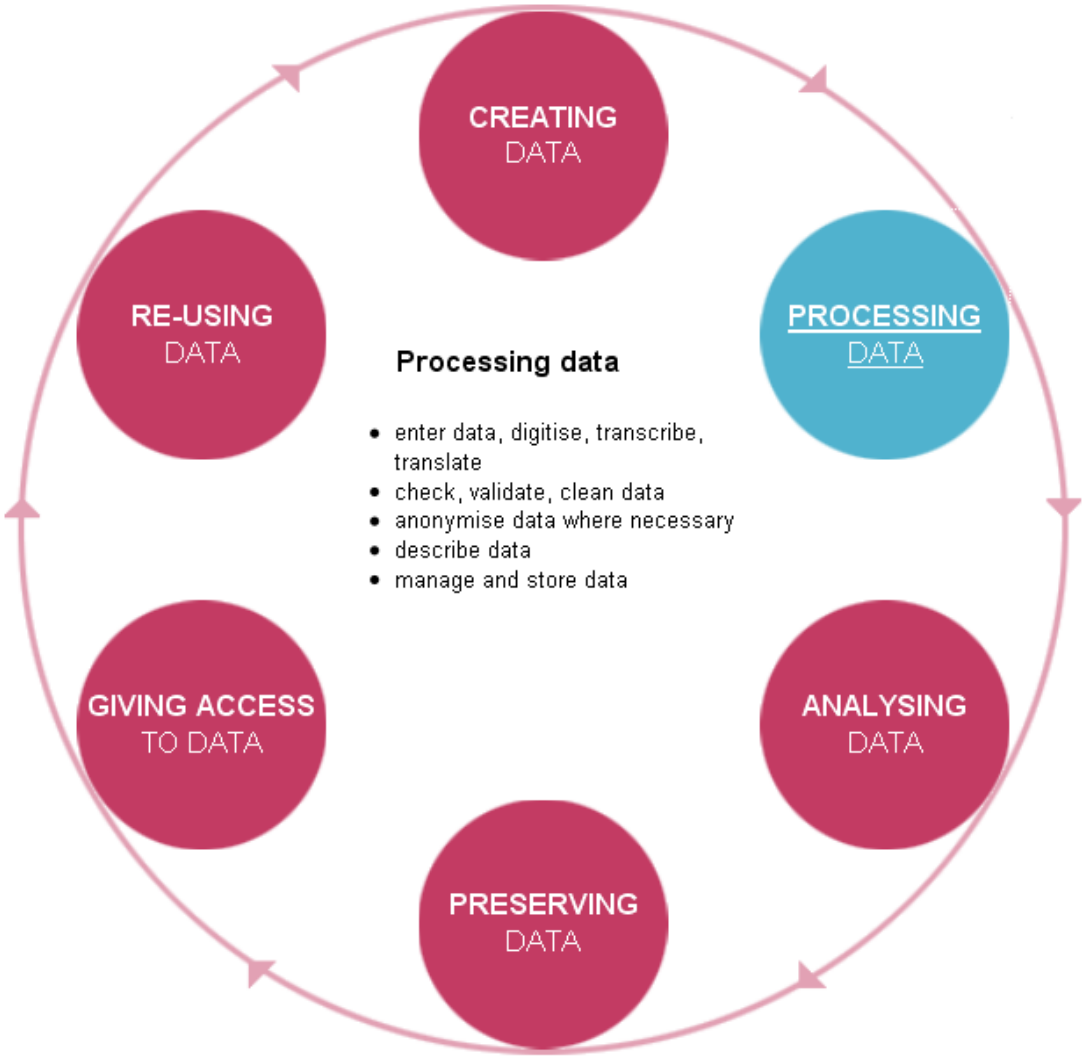
Plan your data management

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

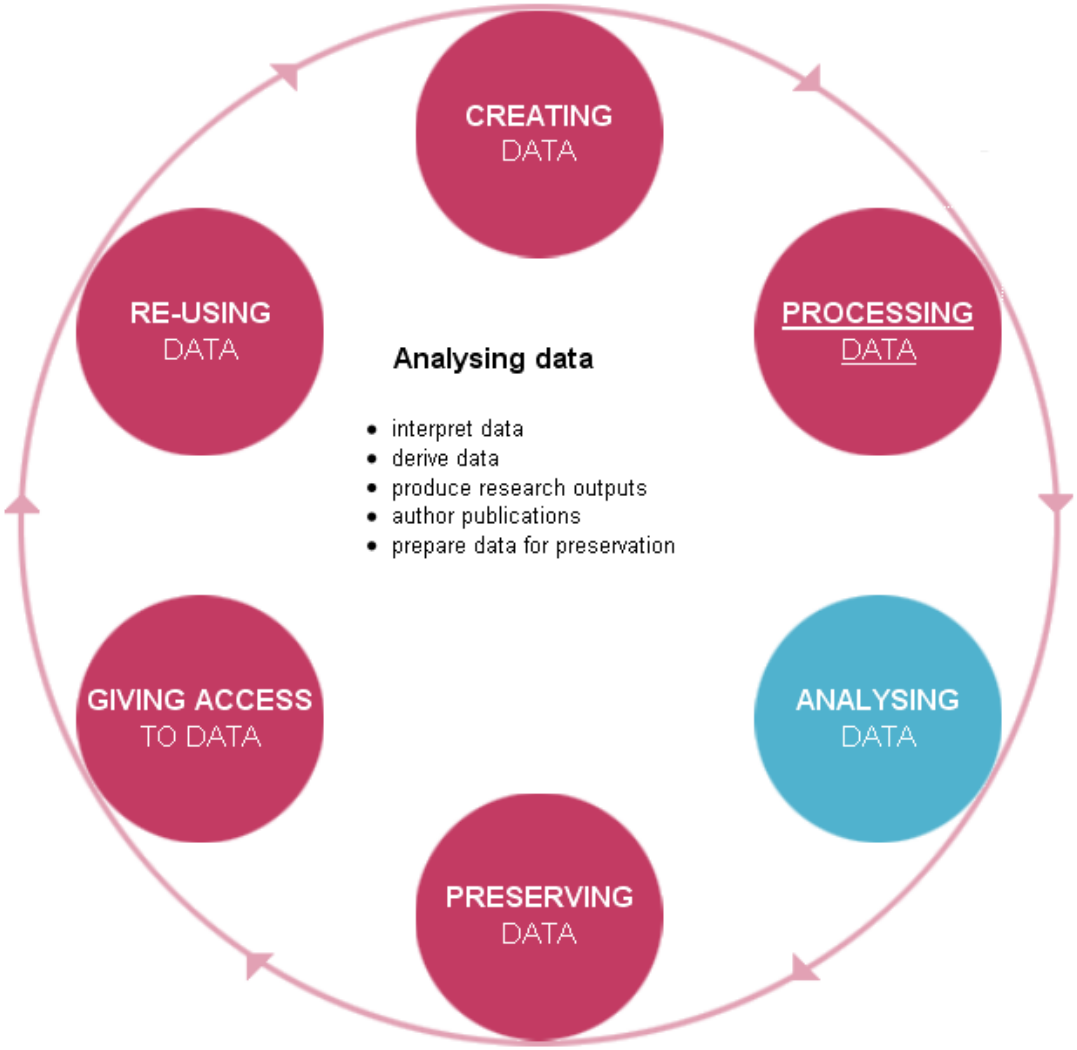
Data life cycle



Data life cycle

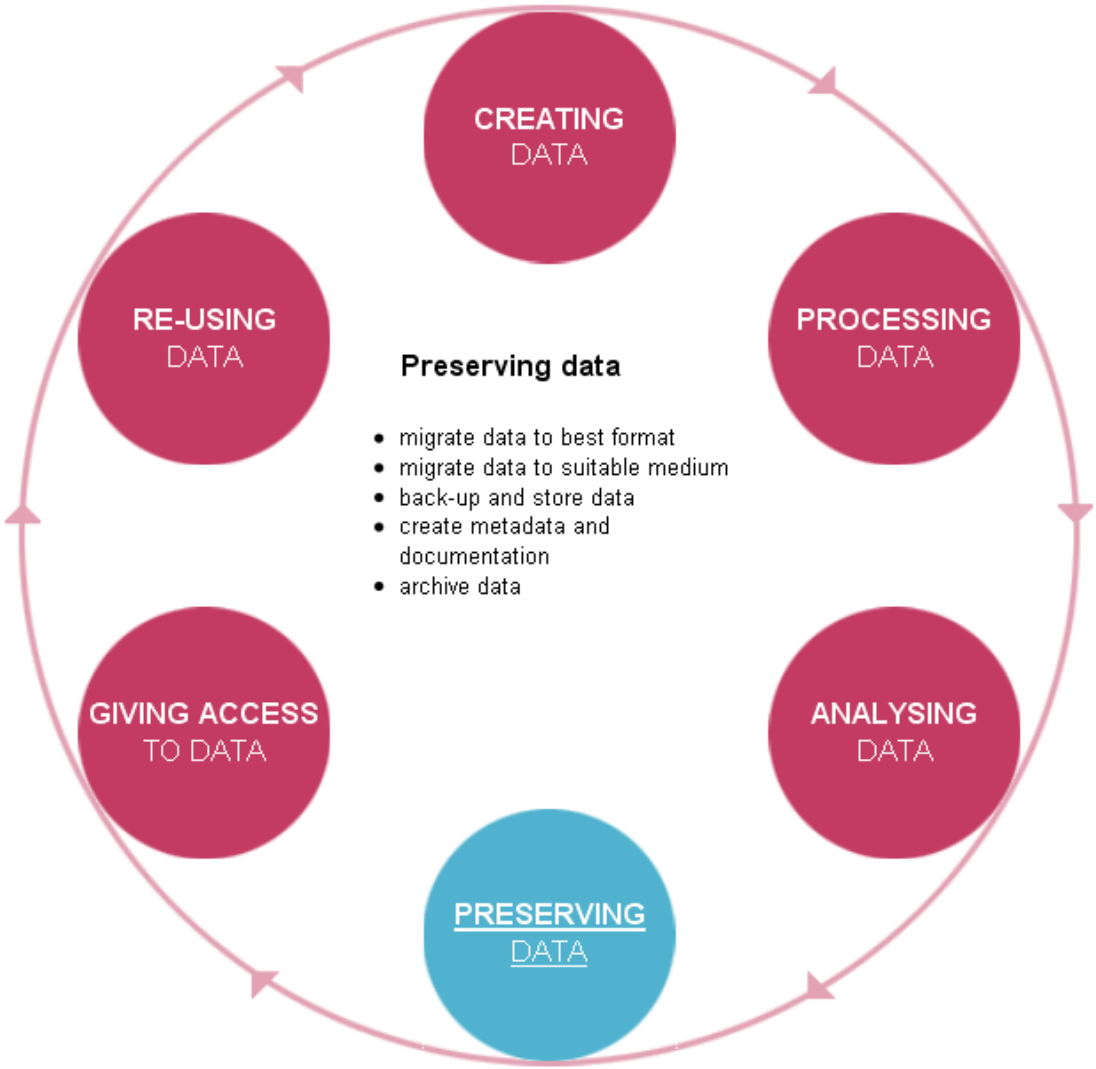


Data life cycle



UK Data Service

Data life cycle



UK Data Service

Data life cycle

UK Data Service

Data life cycle

Exercise: Data Management Checklist

- Have a look through the data management checklist in your binders.

Legal and ethical issues in using and sharing data

Ethical arguments *for* sharing and 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)
- 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

Duty of confidentiality and data sharing

- duty of confidentiality exists in common law and may apply to research data
- an exception to the duty of confidentiality occurs when an informant has consented to information being used in specified ways
- public interest can override duty of confidentiality, so best to avoid vague and general promises in consent forms
  - police investigations
  - child welfare

Data Protection Act, 1998

- personal data:
  - relate to a living individual
  - individual can be identified from those data or from those data and other information
  - include any expression of opinion about the individual
- only disclose personal data if consent given to do so (and if legally required to do do)
- DPA does not apply to anonymised data

- processed fairly and lawfully
- obtained and processed for specified purpose
- adequate, relevant and not excessive for 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

Data protection act and 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

THE *Data Protection Act is not intended to, and does not, inhibit ethical research*

Sensitive data

- data regarding an individual's race or ethnic origin, political opinion, religious beliefs, trade union membership, physical or mental health, sex life, criminal proceedings or convictions (DPA 1998)

- can only be processed for research purposes if:
  - explicit consent (ideally in writing) has been obtained; or
  - medical research by a health professional or equivalent with duty of confidentiality; or
  - analysis of racial/ethnic origins for purpose of equal opportunities monitoring; or
  - in substantial public interest and not causing substantial damage and distress

Options for sharing confidential data

Researchers to consider

- obtaining informed consent, also for data sharing and preservation / curation
- protecting identities e.g. anonymisation, not collecting personal data
- restricting / regulating access where needed (all or part of data) e.g. by group, use, time period
- securely storing personal or sensitive data

Consider jointly and in dialogue with participants

Plan early in research

Do participants consent to share data?

- Timescapes
  - data on personal relationships
  - 95%+ consent rate

- foot and mouth disease in N. Cumbria
  - sensitive community information
  - UK Data Archive consultation; pilot with 4 participants
  - 40/54 interviews; 42/54 diaries; audio restricted

- Finnish research on consent
  - re-contact project: life stores, gender, etc.
  - 165/169 (98%) agreed

- even bereaved relatives want others to benefit from their data

Data sharing and research ethics committees

- RECs are responsible for safeguarding participants from harm and ensuring ethical research (and protecting home institutions), but are not (always) experts in sharing

- there can be perceived tensions between data sharing and protection

- UKDA tries to ensure that RECs know:
  - anonymised data are not subject to DPA
  - many funders require data to be shared
  - most research data can be shared
  - procedures (consent, anonymisation, controlling access) are available to enable ethical sharing
  - data archives ensure ethical re-use of research data, protection of participants and safeguarding of personal data

Exercise: ethical approval forms

Read through the research scenario and imagine yourself as the research on this project.

- What ethical considerations would you need to consider in order to complete this ethical approval form?

- What challenges do you think you might face from the research ethics committee and how would you counter them?

Ethical aspects of protecting and sharing research data: informed consent

UK Data Service

Principles of ethical research

research should be designed, reviewed and undertaken to ensure integrity, quality and transparency

research staff and participants must normally be **informed fully** about the purpose, methods and **intended possible uses of the research**, what their participation in the research entails and what risks, if any, are involved

the **confidentiality** of information supplied by research participants and the anonymity of respondents must be respected

research participants must take part **voluntarily**, free from any coercion

harm to research participants must be avoided in all instances

the independence of research must be clear, and any conflicts of interest or partiality must be explicit.

ESRC Framework for Research Ethics, 2010

A good information sheet and 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
- simple
- avoids excessive warnings
- complete for all purposes: use, publishing, sharing

When to ask for consent

	Pros	Cons
One-off	Simple Least hassle of participant	Research outputs (even questions, not known in advance) Participants will not know all content they will contribute
Process	Most complete for assuring active consent	Might not get consent needed before losing contact Repetitive, can annoy participant

Format for consent

Written

- more solid legal ground (participant agreed to disclose confidential info)
- not possible for some cases: infirm, illiterate, illegal activities
- can be perceived to be off-putting
- can help (or even be ‘required’) by Research Ethics Committees
- may offer more protection for researcher

Verbal - with or without recording

- can be difficult to make all issues clear verbally
- possibly greater risks for researcher
- best if recorded

Law is not specific

- written consent not used on surveys - implicit by taking part

Need to match format of consent with research content

Special cases of participation

Children

- at 16, children can give consent to medical treatments
- under 16s not presumed competent
- but if they ‘understand’, then parental consent not necessary - proceeding without parental approval will be a rare exception

Employees

- issues arise if using workplace for research
- employee may own duty of confidentiality to employer

Vulnerable participants - need to balance

- protection from harm with
- right to participate

Retrospective consent

- covert research, observational psychological experiments (informed??)

Exercise: designing consent forms

Read through the UK Data Archive model consent form.

- How might you adapt this for the purpose of your own research project?

or

- How might you adapt this to fit the community visitor's research project?

Data confidentiality and anonymisation

UK Data Service

Why anonymise research data?

Ethical reasons

- protect people's identity (sensitive, illegal, confidential info)
- disguise research location

Legal reasons

- not disclose personal data (DPA)

Commercial reasons

Discuss with your research participants

Identity disclosure

A person's identity can be disclosed through:

- direct identifiers

  e.g. name, address, postcode, telephone number, voice, picture

often NOT essential research information (administrative)

- indirect identifiers – possible disclosure in combination with other information

  e.g. occupation, geography, unique or exceptional values (outliers) or characteristics

Anonymising quantitative data

- remove direct identifiers
  e.g. names, address, institution, photo

- reduce the precision/detail of a variable through aggregation
  e.g. birth year vs. date of birth, occupational categories, area rather than village

- generalise meaning of detailed text variable
  e.g. occupational expertise

- restrict upper lower ranges of a variable to hide outliers
  e.g. income, age

- combining variables
  e.g. creating non-disclosive rural/urban variable from place variables

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 can distort data, make them unusable, unreliable or misleading

- consistency within research team and throughout project.

- Identify replacements, e.g. with [brackets]

- keep anonymisation log of all replacements, aggregations or removals made – keep separate from anonymised data files

Anonymising qualitative data

Example: Anonymisation log interview transcripts		
Interview / Page	Original	Changed to
Int1		
p1	Spain	European
p1	E-print Ltd	Printing
p2	20th June	June
p2	Amy	Moira
Int2		
p1	Francis	my friend

Other ways of anonymising

- P31. Joan → Mary
- P97. Carol → {Mother}
- P34. Colchester → {Town in S.E.England}
- P65. Welshpool High School → @@##High School##@@

Exercise: anonymising research data

Storing your data: security breaches

UK Data Service

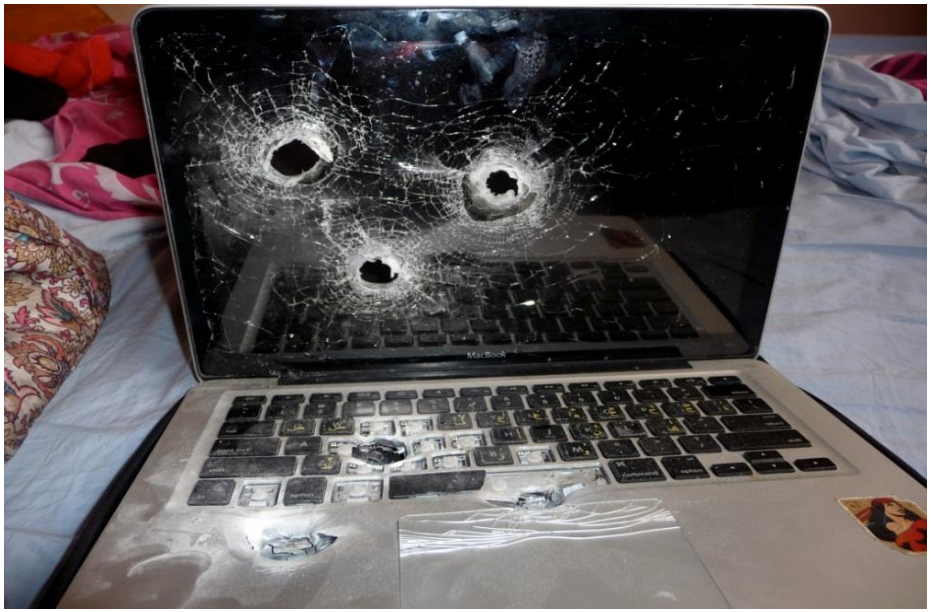
Exercise: security breaches

Think of five possible safety and security breaches which could pose a threat to your data.

1)
2)
3)
4)
5)

Fieldwork from hell

“I’m sorry but we had to blow up your laptop.”



“What….all my client case notes and testimony, writing, pictures, music and applications. Years of work. NO!!!! What?? Are you insane?? What were you thinking? THAT’S ALL MY WORK!?”

Source: [Lilysussman’s Blog]

Data inferno

What if this was your university, your office?

Source: University of Southampton, School of Electronics and Computer Science

Stellenbosch University Fire, South Africa
http://www.youtube.com/watch?v=sFswfvPV4nY&feature=player_embedded#

Backing-up data

- why do back-ups? Risk of loss and change - would your data survive a disaster?
- protect against: software failure, hardware failure, malicious attack, natural disasters
- back-ups are additional copies that can be used to restore originals
- it’s not backed-up unless backed-up with a strategy

Back-up strategy

- consider:
  - what's backed-up? - all, some, just the bits you change?
  - where? - original copy, external local and remote copies
  - what media? - CD, DVD, external hard drive, tape, etc.
  - how often? – assess frequency and automate the process
  - for how long is it kept?
  - verify and recover - never assume, regularly test

- backing-up need not be expensive
  - 1Tb external drives are around £50, with back-up software

Data security

- protect data from unauthorised access, use, change, disclosure and destruction
- personal data need more protection – always keep separate
- control access to computers
  - passwords
  - anti-virus and firewall protection, power surge protection
  - networked vs non-networked PCs
  - all devices: desktops, laptops, memory sticks, mobile devices
  - all locations: work, home, travel
  - restrict access to sensitive materials e.g. consent forms, patient records
- proper disposal of equipment (and data)
  - even reformatting the hard drive is ***not*** sufficient
- control physical access to buildings, rooms, cabinets
- but beware of ‘requirements’ to destroy data

Monday 04 July 2011

The Telegraph

HOME NEWS SPORT FINANCE COMMENT BLOGS CULTURE TRAVEL LIFESTYLI

Technology News | Technology Companies | Technology Reviews | Video Games | Technology

Technology News

Hackers 'steal entire 2011 census'

The entire 2011 UK census database has been stolen by hackers and will be published online, it has been claimed.

Share:

Recommend 9K

Tweet 626

Share 59

+1 0

Technology News

News » Technology » Christopher Williams »

IN TECHNOLOGY

Source: www.telegraph.co.uk

BBC Mobile News | Sport | Weather

NEWS TECHNOLOGY

Home | World | UK | England | N. Ireland | Scotland | Wales | Business | Politics | Health | Educ

8 December 2010 Last updated at 11:43

Nasa sells shuttle PCs without wiping secret data

US space agency Nasa has been left red-faced after selling off computers without ensuring that highly sensitive data had been removed.

An internal investigation found 10 cases where PCs were sold despite failing data removal procedures.

Another four PCs - which were about to be sold - were found to contain data restricted under arms control rules.

Nasa is selling of hundreds of PCs used Space Shuttle programme

Source: www.bbc.co.uk

THE Sun

Monday, 4 July 2011

▶ Log in to comment

HOME
NEWS
Weather
Forces
Sun City
Captain Crunch
Sun Says
Malaria No More
VIDEO
SPORT
Football
Dream Team
Cricket
F1 & Motorsport
SHOWBIZ
Bizarre
Glastonbury
Film
Music
TV
BGT
Soaps
X Factor

News

GOT A STORY? EMAIL : TALKBACK@THE-SUN.CO

EXCLUSIVE

Missing: Laptop with 8.6million medical records

By MIKE SULLIVAN

Published: 15 Jun 2011

Add a comment

A LAPTOP holding the medical records of eight MILLION patients has gone missing.

The computer vanished from an NHS building in the biggest-ever security breach of its kind.

It went missing three weeks ago but has only just been reported to police.

The unencrypted laptop contains sensitive details of 8.63

Missing ... a laptop vanished

TOP STORIES

Monday 04 July 2011

The Telegraph

HOME NEWS SPORT FINANCE COMMENT BLOGS CULTURE TRAVEL LIFESTYLI

Technology News | Technology Companies | Technology Reviews | Video Games | Technology

Technology News

Hackers 'steal entire 2011 census'

The entire 2011 UK census database has been stolen by hackers and will be published online, it has been claimed.

Share:

Recommend 9K

Tweet 626

Share 59

+1 0

Technology News

News » Technology » Christopher Williams »

IN TECHNOLOGY

Source: www.thesun.co.uk

Exercise: security breaches

Read through the following real-life data safety and security breach scenarios.

- Think about what could have been done differently to prevent this happening?
- What preventative measures could the researcher or team have taken?

Our data management services

UK Data Archive Research Data Management Support Services
datasharing@data-archive.ac.uk

Economic and Social Data Services
www.esds.ac.uk/aandp/create

RELU-Data Support service
relu.data-archive.ac.uk

ESRC Research Development Initiative Training Programme
www.data-archive.ac.uk/create-manage/projects/rdi-dm

JISC Data Management Planning project
www.data-archive.ac.uk/create-manage/projects/jisc-dmp

Questions

UK DATA ARCHIVE
UNIVERSITY OF ESSEX
WIVENHOE PARK
COLCHESTER
ESSEX CO4 3SQ
+44 (0)1206 872001

Bethany Morgan Brett
bmorga@essex.ac.uk

Libby Bishop
ebishop@essex.ac.uk