
We recommend you cite the published version.
The publisher’s URL is: http://www.wiserd.ac.uk/resources/wiserd-data-resources/

Refereed: No

(no note)

Disclaimer

UWE has obtained warranties from all depositors as to their title in the material deposited and as to their right to deposit such material.

UWE makes no representation or warranties of commercial utility, title, or fitness for a particular purpose or any other warranty, express or implied in respect of any material deposited.

UWE makes no representation that the use of the materials will not infringe any patent, copyright, trademark or other property or proprietary rights.

UWE accepts no liability for any infringement of intellectual property rights in any material deposited but will remove such material from public view pending investigation in the event of an allegation of any such infringement.

PLEASE SCROLL DOWN FOR TEXT.
Access to sensitive data: Thoughts from the UK experience

Felix Ritchie
The framework principle

- A model of data access

Another model...

user needs → NSI options → solution

legal environment → technology

user needs → principles of access → solution

NSI principles → technology
Law as an enabler

• Up to 2002: various dubious practices
  – £1 contracts
  – Researchers using own equipment
  – Poor records of microdata use

• 2002-2008
  – New recording system for applications
  – Review and rationalisation of legal gateways
  – But still many hurdles to cross

• 2008 –
  – Experience led to significant provision in law for research use
Technology as an enabler

• ‘Spectrum’ of access points balancing
  – value of data
  – ease of use
  – disclosure risk

• for a given level of confidentiality, maximise data use and convenience

• no ‘one-size-fits-all’ solution
  – no absolute prohibitions
  – trade-off is made explicit
  – users determine appropriate level of access
Use of confidential data: the access spectrum

- **Type of access**
  - None
  - VML ONS sites
  - VML Govt sites
  - Secure data service
  - Special licences
  - Licensed data archive
  - Internet

- **Anonymisation**
  - Little
  - Complete

- **SDC of inputs**
  - None
  - Complete

- **Restrictions on users**
  - Many
  - Complete

- **SDC of outputs**
  - Complete
  - None

Distributed access

Distributed data
The framework principle

- user needs
- NSI principles
- principles of access
- technology
- legal environment
- solution
Perceptions of risk and value

• Data access spectrum => all access is doable
  – No more conceptual problems
  – Only question: does risk outweigh value?
    • But if data is already collected, is re-use pure benefit?
    • Does risk assessment include society’s cost not using data?

• Where do risks/benefits arise?
  – Risk of non-use borne by public
  – Benefit of use accrues to public
  – Risk of use borne by NSI?
The framework principle

- user needs
- NSI principles
- legal environment
- principles of access
- technology
- solution
- attitudes to risk
Risk management in practice

• valid statistical purpose

• trusted researchers

• anonymisation of data

• technical controls around data

• disclosure control of results

⇒ safe projects

Active researcher management

safe data

safe setting

Principle-based SDC

⇒ safe use
Active Researcher Management

- Researchers will engage with NSI if given a chance

- Actively engage with researchers
  - In explaining NSI goals
  - In explaining disclosure control
  - In understanding researcher needs, working practices
  - In securing cooperation minimise sensitive output

- Responsibility for data security shared between NSI and researcher (NSI always get final say)

- Certify researchers as part of the security model
Output disclosure control

- Disclosure control at the point of release
- Trained NSI staff and researchers
- Agreement on principles and purpose
- Emphasize co-operation in training
What have we learnt?

- Design based on first principles…
  - made design slow but robust
  - helped identify failings in current approaches
  - showed where new models were needed

- Technology is the easiest problem to solve

- Changes in attitude don’t come easily

- Hindsight is wonderful…
Next stages

• Is the micro/macro distinction still useful?
  – If the ‘how’ is solved, shall we revisit the ‘why’?

• Is short-term vs long-term analysis a more useful distinction?
Questions?

- Felix Ritchie
- Microdata Analysis and User Support
- felix.ritchie@ons.gsi.gov.uk

- Virtual Microdata Laboratory (VML)
- Microdata Analysis and User Support
- maus@ons.gsi.gov.uk