# Thinking beyond the Tick Box

## complying with the spirit and not just the letter of the law

Andrew Charlesworth
Centre for IT & Law,
University of Bristol

# Is it me or is it warmer....?

- Research data can be of great interest to 3rd parties.
- Where the research is carried out within a public authority, the FOIA 2000 may apply to it.
- Failure to comply with FOI requests, in whole or in part, may expose institutions to regulatory or legal action.
- It may also result in bad publicity for an institution, particularly where the data concerned is high-profile.
- May compromise or tarnish otherwise successful and high value research.

# In the hot seat

- What was it about the UEA climate change data that made it controversial?
  - Could the problems have been foreseen by the researchers or by their institution?
  - Climate change is a 'hot button' topic – high interest to journalists, climate change deniers, environmentalists
  - In retrospect, it is unsurprising that there was demand for access to the data - in the 'public interest'
  - Equally unsurprising that scientists are not keen on other people 'imaginatively re-interpreting' their data
  - How to tackle the risks? Institutional role?

# Storing up Trouble

- Faculty Research Committee edict – all research data must be held for x period of time
- Problem – who is responsible for curating that data?
- Past experience suggests that leaving safe storage to researchers over long periods of time is problematic
  - Staff turnover and equipment replacement
  - Advances in technology/software
  - Inadequate security for personal data (DP, confidentiality)
  - Ability of institution to locate data at short notice (FOIA)
  - Ability of institution to determine when data should finally be weeded/deleted, and who should do it (ownership of data)
  - Loss of surrounding data, loss of context

# Playing gatekeeper

- Imposing structure on current processes
  - Ethics Committee – tying in ethical approval with effective data management – departmental data deposit and audit.
  - Process requires a range of descriptive information, but also a 'start and end' risk assessment
    - what was going to be collected, what was collected,
    - what risks were envisaged, what risks have developed
    - what's the rationale for storage over x period of time
    - what's the likelihood of internal/external access during that time – schedule for curation

# Problems

- Most legal and ethical problems arise because of:
  - Lack of effective control (ownership/guardianship)
  - Lack of appropriate/accurate information
  - Poor understanding of legal and ethical issues by researchers (or refusal to engage, or deliberate misinterpretation – 'confidential')
  - Failure to adjust policies and practices to new circumstances (law, technology, politics - evergreening)
  - Lack of sanction (where do consequences of data loss, data breach, data misuse fall?)

# Solutions?

- Standard solution to data-related risk is often ‘a policy.’
- All researchers are supposed to follow ‘the policy’.
- To help them comply there is usually ‘a form’.
- Problems:
  - Policies are rarely ‘one size fits all’
  - Policies are often not ‘evergreened’
  - Most forms are effectively a checklist
  - Checklists tend to breed complacency (checklist mentality)
  - Complacency is fatal to effective risk assessment and risk management

# Thinking outside the tick box

- To effectively curate datasets, metadata collection will be necessary – there is a place for forms
- BUT we need people to think clearly about risks in the data they are collecting or archiving
- It may be possible to learn from developments in DP law – privacy impact assessments, privacy by design.
  - How do we determine risk in particular data/datasets?
  - Is the context of collection and storage a factor?
  - Are there special factors in play – ethical issues?
  - What issues should determine length of storage, security of storage, degree of gatekeeper control etc?