Encouraging access to confidential research data

Presentation at the 40th IASSIST Conference
Toronto, June 3-6, 2014

Dr. Katharina Kinder-Kurlanda
GESIS - Leibniz Institute for the Social Sciences
katharina.kinder-kurlanda@gesis.org

Laurence Horton
The London School of Economics and Political Science
@laurencedata, L.Horton@lse.ac.uk
Legal context in the UK

- Data Protection Act (DPA) (1998)
- Information Commissioner Office (ICO) guidance and commissioner precedents
Legal context in Germany

- Federal Data Protection Act = Bundesdatenschutzgesetz (BDSG)
  “The purpose of this Act is to protect individuals against infringement of their right to privacy as the result of the handling of their personal data. “
  “Personal data” = any information concerning the personal or material circumstances of an identified or identifiable natural person
Anonymisation in the UK

"Anonymisation is the process of turning data into a form which does not identify individuals and where identification is not likely to take place."

See: http://www.ico.org.uk/for_organisations/data_protection/topic_guides/anonymisation
Anonymisation in Germany

- BDSG, §3,6 on 'factual anonymisation':
  “Rendering anonymous’ shall mean the alteration of personal data so that information concerning personal or material circumstances cannot be attributed to an identified or identifiable natural person or that such attribution would require a disproportionate amount of time, expense and effort.”
Legal context of the EU

- 1995 Data Protection Directive
- Personal data is any information relating to an individual
- Apply when a person can be identified, directly or indirectly
- Implemented (unevenly) through member states
Current issues for researchers

Lack of harmonization in data access regimes due to legal uncertainty
• What's fine for a Belgian is not fine for the Dutch
• Gaps, inconsistencies and contradictions
• Under what circumstances are small cell values problematic...?
• What degree of security measures is 'enough'?
• Does the door of a safe room need to be closed...?
Current issues for researchers (2)

Technological developments are ahead of legal ones

• Therefore 'best practice' has become very important as everybody is cautious (Avoiding unnecessarily conservative approaches)
  – e.g.: the proposed EC current reforms
Current issues for researchers (3)

Balancing researchers' needs against a public mood of (understandable) suspicion

- Example: high profile cases of personal data leakage (not research data, but still...)
- Example: Albrecht report: Legitimate concerns regarding marketing interests pushing for cutting down on privacy lead to researchers being 'punished'
  - But: Data becomes more transparent through research!
Archives to the rescue?

• Data archives are expanding their role within data infrastructures. Not just end point data providers, but support services for those creating data
  – Interest in ensuring good quality data is produced
Archives to the rescue?

• Archives build up a store of expertise through experience
  – Encounter a lot more data and a diversity of projects that most researchers never encounter
Where archives help

• Combating ignorance in the research community about data protection laws, by...
• Active help with specific queries
• Proactive help: Creating training and support materials
• Providing the technological capacity to share data that could not ordinarily be shared
• Legally binding user licenses
• Secure data services for rich but sensitive data
Example: GESIS SDC - Secure Data Center

- Contractual safeguards
- Technical safeguards
- Organisational safeguards
Aims of the GESIS Secure Data Center

To provide controlled and secure access to data deserving special protection

Secure Data Center (SDC)

- Off-site access with a legal contract
  - Mostly anonymised

- On-site access in the SDC Safe Room
  - Some anonymisation

- Remote access (in the planning stage)
  - Some anonymisation
GESIS Secure Data Center off-site access

- Users sign a contract and agree
  - …not to attempt identification of individuals
  - …not to share the data with anyone
  - …to keep the data secure
GESIS Secure Data Center Safe Room

- Users sign a contract and access the data in a secure and controlled environment:
- Access control, no mobile phones, no laptops…
- Thin client with a virtual machine at GESIS
Controlled data access within institutions

LSE Library
• Manage applications for and use of sensitive data
  • Examples
    • UK Data Service applications
    • Eurostat EU-Statistics on Income and Living Conditions (EU-SILC)
    • Household, Income and Labour Dynamics in Australia (HILDA)
Controlled data access within institutions

Future developments

• Expectation to improve data security procedures and technology
  • Possible establishing university safe data rooms
  • Working closely with IT providers on data security and meeting data provider requirements
Thank you