Course for Doctoral Students

RESEARCH DATA MANAGEMENT AND OPEN DATA

23rd July 2015, Social Science Data Archives, Faculty of Social Sciences, University of Ljubljana

ECPR Summer School 2015
LEGAL AND ETHICAL ISSUES: 
CONSENT, ANONYMISATION, 
ACCESS REGULATION 

Veerle Van den Eynden, UK Data Service
Options for sharing research data that may contain confidential information

- 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
Consent across the data life cycle

- Engagement in the research process
- 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 research data - for primary research and sharing
  • strategies to ensure confidentiality of data (anonymisation, access etc.) where this is relevant

• Need to balance
  • as simple as possible
  • complete for all purposes: use, publishing, sharing
  • avoid excessive warnings
## Timing of consent

<table>
<thead>
<tr>
<th></th>
<th>PROS</th>
<th>CONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-off</td>
<td>• Simple</td>
<td>• Research outputs not known in advance</td>
</tr>
<tr>
<td></td>
<td>• Least hassle to participants</td>
<td>• Participants will not know all info they will contribute</td>
</tr>
<tr>
<td>Process</td>
<td>• Ensures ‘active’ consent</td>
<td>• May not get all consent needed before losing contact</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Repetitive, can annoy participants</td>
</tr>
</tbody>
</table>
## Form of consent

| Written | • More solid legal ground, e.g. participant has agreed to disclose confidential info  
| | • Often required by IRB  
| | • Offers more protection for researcher  
| | • Not possible for some cases: infirm, illegal activities |
| Verbal | • Can be difficult to make all issues clear verbally  
| | • Possibly greater risks for researcher  
| | • Best if recorded |
Aspects to consider

• Different forms of consent for different materials, e.g. audio recordings vs transcripts
• Right to withdraw - what to do with already collected data?
• Informed consent for ‘unknown future data uses’?
• Provide maximum information about reuse
  • who can access the data - authenticated researchers
  • purposes - research or teaching or both
  • confidentiality protections; agreement by future users
Medical research and biobanks

- Enduring, broad, open consent
- No time limits; no recontact required
- Unspecified hypotheses and procedures
- 99% consent rate (2500+ patients) - Wales Cancer Bank
Types of material and consent

Different data sharing consent agreements may be applied to different types of research data, e.g. less sensitive (survey) vs. highly sensitive (medical)

- Text and transcripts
  - can be anonymised
- Images, audio/video recordings
  - data more likely to reveal identities
  - less usable after anonymizing (distortion or blurring)
  - anonymising costly

Consent or access control may be better alternatives than anonymisation
Special cases of consent

Children
  • own consent (>16) or parent/guardian consent

Employees
  • duty of confidentiality to employer, e.g. employment-related research

Vulnerable people, disabilities of any kind
  • balance protection from harm with right to participate in research

Internet research, blogs, social media - public vs. private information, can consent be asked?
  • ethical decision-making = deliberative process
  • nsrnss.blogspot.co.uk/2014/02/new-social-media-new-social-science-and.html
  • aoir.org/reports/ethics2.pdf

Retrospective consent, covert research, observational experiments
Do participants consent to share data with other researchers in future?

- Foot and mouth disease in N. Cumbria
  - sensitive community information
  - 40/54 interviews; 42/54 diaries; audio restricted

- Finnish research on consent
  - re-contact participants: life stories, gender, etc.
  - 165/169 (98%) agreed to sharing

- Timescapes
  - longitudinal data on personal relationships
  - > 95% consent rate to share data
In practice: wording in consent form / information sheet

We expect to use your contributed information in various outputs, including a report and content for a website. Extracts of interviews and some photographs may both be used. We will get your permission before using a quote from you or a photograph of you.

After the project has ended, we intend to archive the interviews at .... Then the interview data can be disseminated for reuse by other researchers, for research and learning purposes.

The interviews will be archived at ....... and disseminated so other researchers can reuse this information for research and learning purposes:

- I agree for the audio recording of my interview to be archived and disseminated for reuse
- I agree for the transcript of my interview to be archived and disseminated for reuse
- I agree for any photographs of me taken during interview to be archived and disseminated for reuse
In practice: wording in consent form / information sheet

<table>
<thead>
<tr>
<th>Use of the information I provide beyond this project</th>
<th>Y</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>I agree for the data I provide to be archived at the UK Data Archive.</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>I understand that other genuine researchers will have access to this data only if they agree to preserve the confidentiality of the information as requested in this form.</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>I understand that other genuine researchers may use my words in publications, reports, web pages, and other research outputs, only if they agree to preserve the confidentiality of the information as requested in this form.</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

As the ESRC is a publicly funded body, it has developed ways to share data among academic researchers (subject to strict conditions). To this end, we hope you will allow your anonymised transcript to be stored as part of the UK Data Archive (a service provider for the Economic and Social Data Service).
In practice: wording in consent form / information sheet

Any personal information that could identify you will be removed or changed before files are shared with other researchers or results are made public.

UK Data Archive model consent form
[link: ukdataservice.ac.uk/manage-data/legal-ethical/consent-data-sharing/consent-forms.aspx]
Anonymising research data

- Direct identifiers - often not essential research info
- Indirect identifiers

- Remove direct identifiers (or replace with pseudonyms)
  e.g. names, address, institution, photo

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

- Generalise meaning of detailed text
  e.g. occupational expertise

- Restrict upper lower ranges to hide outliers
  e.g. income, age
Anonymising qualitative data

- Remove direct identifiers, or replace with pseudonyms - often not essential research info
- Avoid blanking out; use pseudonyms or replacements
- Identify replacements, e.g. with [brackets]
- Plan or apply editing at time of transcription
- Avoid over-anonymising - removing information in text can distort data, make them unusable, unreliable or misleading; so balance anonymisation with the need to preserve context
- Consistency within research team and throughout project.
- Keep anonymisation log of replacements or removals made - keep separate from anonymised data files
In practice: example anonymisation

Ex 1. Health and Social Consequences of the Foot and Mouth Disease Epidemic in North Cumbria, 2001-2003 (study 5407 in UK Data Archive collection) by M. Mort, Lancaster University, Institute for Health Research.

Date of Interview: 21/02/02

Interview with [Lucas Roberts], DEFRA field officer
Date of birth: 2 May 1965
Gender: Male
Occupation: Frontline worker
Location: Plumpton, North Cumbria

[Lucas] was living at home with his parents, "but I'm hoping to move out soon" so we met at his parents' small neat house. We sat in a very comfortable sitting room with an open fire and [Lucas] made me coffee and offered shortbread. Although at first [Lucas] seemed a little nervous, quick to speech and very watchful he seemed to relax as we spoke and to forget about the tape.

I will just start by asking you to tell me a little bit about yourself and your background.

Well it is an agricultural background. I grew up on the farm where my brother is now. After I left school I did work on the farm but went to college and did exams, did land use recreation, sort of countryside/environmental management course. So I obviously left agriculture, did the course and came back [to the farm] at weekends.
In practice: example anonymisation

Yeah. So is part of your job to look for funding bids and to write funding bits or is that separate?

No. That was what P3 used to do and then it sort of passed down to... really it's with P1 and P4. But I don't actually think there's anything out there at the moment. I think at the moment, because there's all this money saving and things, there's nothing to... there isn't actually anything to access.

M27UK

INT: Id like to start with your career history really. Perhaps you can start from when you finished school?

RES: First of all Id like to say that I do not think that my career is very traditional for someone who comes and works abroad. I graduated from mathematics in ‘Country F’. That was equivalent to what is now BSc and MSc and then I did a separate MSc in Computer Science which at that time was called like Specialisation in computer science. And then I started working at the ‘Country F’ University as a teacher.

INT: Was that in ‘City D’?

RES: No, my Specialisation in Computer Science was in ‘City D’ and my first degree, my mathematics degree was in another university in ‘Country F' it is called [Text Cut]

So after I finished in ‘City D’ my specialisation in computer science I went back to [Text Cut] where I did my first degree and I started working there as an assistant Professor.
Managing access to data

**Open**
- available for download/online access under open licence without any registration

**Safeguarded**
- available for download/online access to logged-in users who have registered and agreed to an End User Licence

**Controlled**
- available for remote or safe room access to authorised and authenticated users whose research proposal has been approved and who have received training
Can such research data be open?

• ESRC research data policy:
  • *Publicly-funded research data are a public good, produced in the public interest, which shall be made openly available and accessible with as few restrictions as possible in a timely and responsible manner that meets a high ethical standard and does not violate privacy or harm intellectual property.*

  • Openly available research data, with as few restrictions as possible, means in the ESRC context that *research data will be made available for re-use free of charge, as open data, safeguarded data or controlled data; the access category being selected to minimise the risk of disclosing personal information*
Open about data with restricted access

Publish:
• Which data exist
• Where data are kept, e.g. which repository
• Who can access them
• For which purpose can they be used
• Under which conditions
In practice: data with access conditions

Health and Social Consequences of the Foot and Mouth Disease Epidemic in North Cumbria, 2001-2003 (study 5407 in UK Data Archive collection) by M. Mort, Lancaster University, Institute for Health Research.

- Interviews (audio + transcript) and written diaries with 54 people
- 40 interview and diary transcripts are archived and available for re-use by registered users
- 3 interviews and 5 diaries are embargoed until 2015
- audio files archived and only available by permission from researchers

[discover.ukdataservice.ac.uk/catalogue/?sn=5407](discover.ukdataservice.ac.uk/catalogue/?sn=5407)
[doc.ukdataservice.ac.uk/doc/5407/mrdoc/pdf/q5407userguide.pdf](doc.ukdataservice.ac.uk/doc/5407/mrdoc/pdf/q5407userguide.pdf)
In practice: access conditions ReShare

Global Uncertainties: Security In an Africa of Networked, Multi-Level Governance


The programme of research centres on how the various institutions responsible for the production of security and the management of conflict in Sub-Saharan African societies do, could and should evolve in response to the presence of violent conflict. These institutions include: the armed forces, police, courts (civil as well as criminal), 'traditional' institutions of local governance, elections (and other methods of choosing leadership), international organisations (such as the United Nations, African Union, UN High Commission for Refugees, the World Court), and the community of international donor nations. The programme is built on the observation that all governance (especially in Africa) is multi-leveled and networked - from the village to the international organisation, and well beyond what is specified in formal government structures. Thus the focus will be not only on the ways in which key conflict-management institutions evolve themselves but also on the changing ways in which the networks in which they are embedded actually operate. This leading edge research challenges theories about the state, state formation, and sovereignty as well as the dynamics of violent conflict as presented in the dominant current work on Africa. The primary methods of research will be network analysis and qualitative interviewing.

Creators: Leonard, David Email: Unspecified
Affiliation: Institute of Development Studies ORCID: Unspecified

Research funders: Economic and Social Research Council
Grant reference: RES-071-27-0048
Subjects: Law, crime and legal systems, Politics
Date deposited: 31 Jan 2014 19:00
Last modified: 22 Oct 2014 19:27
In practice: access conditions ReShare

Coverage and Methodology

Collection period:
Date from: 1 April 2009
Date to: 31 March 2013
Country:
Sierra Leone
Data collection method:
Group interviews in 39 villages and towns in rural Sierra Leone
Observation unit:
Groups
Kind of data:
Alpha-numeric, Textual
Type of data:
Qualitative and mixed methods data
Resource language:
English

Access and Administration

Data sourcing, processing and preparation:
Citizen interviews
Copyright holders:
Leonard, David
Email: Unspecified
Affiliation: Institute of Development Studies
ORCID: Unspecified
Leonard, David
Email: Unspecified
Affiliation: Institute of Development Studies
ORCID: Unspecified

Notes on access:
The names of individual respondents have been removed from the data set.
Publisher:
Economic and Social Research Council
Last modified:
22 Oct 2014 19:27
### AVAILABLE FILES

#### Data

<table>
<thead>
<tr>
<th>File or bundle content:</th>
<th>Security &amp; Networks</th>
</tr>
</thead>
<tbody>
<tr>
<td>File format:</td>
<td>application/octet-stream</td>
</tr>
<tr>
<td>License:</td>
<td>UK Data Service End User Licence</td>
</tr>
<tr>
<td>File size:</td>
<td>10Kb</td>
</tr>
</tbody>
</table>

- [Security_%26_Networks.xlsx](#) **Registered users only (safeguarded data)**

- [Sierra_Leone_Security_%26_Networks__Coded.xls](#)

- [Copy_of_Somalia_SC_peace_initiatives.xlsx](#)

#### Documentation

<table>
<thead>
<tr>
<th>File or bundle content:</th>
<th>Sierra Leone methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>File format:</td>
<td>application/msword</td>
</tr>
<tr>
<td>License:</td>
<td>UK Data Service End User Licence</td>
</tr>
<tr>
<td>File size:</td>
<td>64Kb</td>
</tr>
</tbody>
</table>

- [Sierra_Leone_methods.doc](#) **Anyone (open data)**
In practice: access conditions ReShare

The Impact of Living Abroad: Stress, Adaptation, and Intergroup Contact


Whether as migrant, expatriate, or overseas student, living abroad has an impact on individuals in terms of well-being, task-oriented functioning, and intergroup behaviour. This research examines the impact of intercultural contact by comparing such sojourners with a matched group of peers. The study will monitor a large cohort of international exchange students (sojourners) before, during and after intercultural contact. The research has four objectives. A first objective is to research the occurrence of acculturative stress, specifically examining when, and why stress occurs, and how it evolves over time. A second objective is to examine cultural learning and adaptation. In particular, the research will explore whether sojourners benefit in terms of improved culture specific knowledge (e.g. language, customs), general cultural awareness, and personal development (e.g. self-confidence). A third objective is to study the impact of Intercultural contact on intergroup behaviour. The study will appraise whether and how intergroup contact influences perception of ingroup and outgroups. Finally, the study will investigate how differences in cultural distance (i.e. the extent to which cultures are different) affect these issues of stress, adaptation, and intergroup contact.

Creators:
Geeraert, Nicolas
Unspecified

Research funders:
Economic and Social Research Council

Grant reference:
RES-062-23-1211

Subjects:
Psychology

Data deposited:
04 Apr 2013 13:42

Last modified:
01 Oct 2014 13:22

Available files:

Data
- data.csv

Accessible to:
Registered users requesting permission access from depositor until 1 January 2015

File or bundle content:
Data

File or bundle description:
data
Questions?

- Veerle Van den Eynden
- veerle@essex.ac.uk