Ethics Studio

Identifiability of information and data management.

Dianne Vella-Brodrick and Tim Mattingsbrooke
21 November, 2018

NH&MRC National Statement on Ethical Conduct in Human Research – Section 3, Chapter 3.1, Element 4.
NH&MRC National Statement on Ethical Conduct in Human Research

- ‘data’ - bits of information in their raw form,
- ‘information’ - data that have been interpreted, analysed or contextualised.
Examples of data and information include:

- What people say (e.g., interviews, surveys, biographies)
- Images and audio visual material
- Administrative records (billing, service details, health notifications)
- Digital information (mobile devices and on-line)
- Physical specimens or artefacts and biospecimens (blood, bone, muscle, urine)
- Analysis of existing personal information (clinical, health records)
- Observations
- Results from experiments
Data and information

- Generation
- Collection
- Access
- Use
- Analysis
- Disclosure
- Storage
- Retention
- Disposal
- Sharing
- Re-use
Key questions

NH&MRC National Statement on Ethical Conduct in Human Research – Chapter 4

Collection:
• What data or information are required to achieve the objectives of the project?
• How and by whom will the data or information be generated, collected and/or accessed?

Use:
• How and by whom will the data or information be used and analysed?
• Will the data or information be disclosed or shared?

Management:
• How will the data or information be stored and disposed of?
Identifiability of information and risk

Higher identifiability of information typically leads to great risk of harm or discomfort.

Examples of identifiers:
- Name
- Image
- Date of birth
- Address
- Attribute
- Group affiliation (Principal at X School)

De-identification (e.g., use of code) and re-identification
Personal information

Ensure that you:

- Collect the minimum possible amount of personal information.
- If possible, collect the necessary data without using personally identifying information.
- If personally identifying information is required, de-identify your data upon collection, or as soon as possible thereafter, and store identification keys in a separate, secure location.
- Separate the roles of those responsible for managing identifiers and those responsible for managing content.

Personal information classed as ‘sensitive information’, is subject to a higher level of privacy protection than other ‘personal information’.
Confidentiality

- We intend to protect the confidentiality of your responses to the fullest possible extent, within the limits of the law (i.e., we will disclose information about your to the school only if we feel that your or other people’s health and safety is seriously at risk). Your name and contact details will be kept in a separate, password-protected computer file from any data that is provided. No identifying details are required on survey responses. However, information provided as part of group discussions will not be anonymous although the importance of keeping information confidential will be explained to all group participants who will be advised to share only information that they are comfortable sharing. Also, only group data will be reported, and therefore your individual responses will remain confidential in any public documents or presentations.
Privacy at the University of Melbourne

- The University has obligations regarding the collection, use and management of personal and health information under various pieces of legislation, including the *Privacy and Data Protection Act 2014 (Vic)* and *Health Records Act 2001 (Vic).*

- The University’s *Privacy Policy (MPF1104)* further governs privacy at the University. This policy applies to all areas of the University. Any person who deals with the University is entitled to expect that their personal or health information will be treated in accordance with this policy.

- The University Secretary is the University’s designated Privacy Officer. The Privacy Coordinator within Legal and Risk, provides guidance on privacy obligations and responsibilities.
At UoM data must always be stored for at least five years after any publication/public release has arisen from that data.
  - digital or paper records
  - held directly by a senior academic or by the institution.

These data and records should normally be shared, upon request, with other researchers, with reference to funding agreements and privacy considerations.
  - Melbourne.figshare or discipline-specific repositories.

Research data made available for re-use and public scrutiny
  - maximises transparency and accountability
  - provides safeguards against misconduct, such as data fabrication and data falsification.
5.3 Data Management

How do you propose to manage the data collected in this research project? Specify what types of data will be collected, how they will be stored and in what format. How will access to the data be controlled and by whom? Discuss retention, security, and data sharing plans. What measures will be taken to protect participants’ privacy, and their data?

Refer to NS §1.11, the Australian Code for Responsible Conduct of Research §2, and the University of Melbourne Code of Conduct for Research §2.1.

A) Privacy and Confidentiality

Data will be stored electronically in a secured hosting database and will be carefully managed with unique ID codes to ensure that participant responses cannot be identified by anyone other than the research team (Vella-Brodrick, co-researchers, associated personnel, research students). Nobody except the responsible researcher and trusted personnel will have access to lists that match participants’ actual names against their ID codes. Names of participants on signed consent forms will be securely stored separately to the rest of the data. Interview data will be audio-recorded and participants will be asked to avoid using personal identifiers. No identifying details of participants will be published or made publicly available.

B) Security and Storage of Data

Vella-Brodrick will be responsible for managing the security and storage of data, and only Vella-Brodrick and her trusted personnel will have password-protected access to the raw data. Records will be stored in conditions suitable to the nature of the record content (e.g., personal, sensitive or confidential information), that is, in securely locked facilities at the Melbourne Graduate School of Education through which the project is being conducted.

C) Retention

Research data will be retained for a minimum retention period of five years after the last publication or public release arising from the research.
Managing Data @ Melbourne

Responsibilities relating to data include:

• making and keeping complete, clear, and accurate records of all research;
• sharing findings and data openly and promptly, as soon as you have established priority and ownership claims;
• publishing and communicating research honestly and accurately;
• employing appropriate methods and using a high level of rigour and objectivity in research activities;
• appropriately citing and, where applicable, obtaining permission for the use of all published and unpublished work;
• acknowledging in research outputs all contributors and contributions to the research data described in the research output.
Nicola's file management

In the last section, Nicola discussed the file naming system used by her research group. Here is a screenshot showing that file naming system in more detail. You can see that each file begins with a four-letter project code, then the date (in a consistent format), the title of the file, and finally the extension. The advantages to using this system are many. Using a consistent project code means that the file can be linked to a project even if it is copied to another location. Note too how the documents are organised into sub-folders. this has been done in a way that makes sense for this research group - a different way of organising might make sense in your field of research.

<table>
<thead>
<tr>
<th>Name</th>
<th>Date Modified</th>
<th>Size</th>
<th>Kind</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCAG</td>
<td>Today, 4:09 pm</td>
<td>30.6 MB</td>
<td>Folder</td>
</tr>
<tr>
<td>docs 2016 to</td>
<td>Today, 4:09 pm</td>
<td>30.6 MB</td>
<td>Folder</td>
</tr>
<tr>
<td>Part 1 2016-07-20 to</td>
<td>11/1/17, 3:08 pm</td>
<td>30.6 MB</td>
<td>Folder</td>
</tr>
<tr>
<td>MCAG - 2016-07-22 Email from Rimi re project report and Tia.pdf</td>
<td>22/7/16, 3:48 pm</td>
<td>40 KB</td>
<td>PDF Document</td>
</tr>
<tr>
<td>MCAG - 2016-07-22 M&amp;G Year 4 report.pdf</td>
<td>22/7/16, 3:47 pm</td>
<td>10 MB</td>
<td>PDF Document</td>
</tr>
<tr>
<td>MCAG - 2016-11-23 Email from Rimi re website.pdf</td>
<td>24/11/16, 1:46 pm</td>
<td>91 KB</td>
<td>PDF Document</td>
</tr>
<tr>
<td>MCAG - 2016-11-23 Website mockup 2.pptx</td>
<td>24/11/16, 1:43 pm</td>
<td>20.5 MB</td>
<td>PowerPoi...tion (.pptx)</td>
</tr>
<tr>
<td>Visualisations 2016 to</td>
<td>21/11/16, 12:02 pm</td>
<td>29 KB</td>
<td>Folder</td>
</tr>
</tbody>
</table>
Security and Storage

Security and Storage of Digital Data:

- For the duration of the study, all electronic files including de-identified data spreadsheet, will be stored on a password protected computer, these will be backed up to figshare. UoM figshare service provides a repository of raw data, accessible for analysis and publication and available for supervisor access. Figshare is suitable for long term data storage. All data will be destroyed after 5 years from the last publication or public release.
• A scholar from the UK who works in a similar area to you has contacted you to ask you for your data. She has replicated your study and has found very different results from you. What do you do? What factors do you need to consider?
• You have collected information for one purpose but then realized that this data can also be used for a different (secondary) purpose.

• For example, you have collected data on conceptualisations of success with young athletes. In a separate study you have also collected conceptualisations of success from UAE students. You now want to do a cross-country comparison.

• What are some of the factors that will need to be considered before this data can be ‘re-used’?
Thank you

Next Ethics Studio:
• 10 Dec