Oral history and personal data: recording, archiving and re-use under GDPR

Rob Perks, Oral History, British Library
www.bl.uk/oralhistory
www.ohs.org.uk/advice/data-protection
Key legal and ethical issues for oral history

• ‘Consent’, permission (university ethics committees)
• Copyright (1988 but with 2014 exceptions)
• Creative Commons (= licensing)
• Libel
• Illegal activities, criminal investigation (eg Boston College)
• Freedom of Information (FoI)
• Safeguarding (policies when working on certain projects)
• Data Protection (GDPR)
Data Protection

• General Data Protection Regulation (GDPR) came into effect in the UK from 25 May 2018, alongside the new UK Data Protection Act 2018.

• When/if we leave the EU a new UK GDPR will be implemented, not expected to be much different but it will only apply to UK data not to EU data (= implications for joint UK/EU projects sharing data)

• GDPR applies to any organisation, individual or group which collects personal data, including oral history projects, charities, community organisations, youth groups, libraries, museums, archives, educational organisations etc but also individual researchers, whether salaried, self-employed or voluntary.
What is personal data?

- GDPR gives individuals certain **rights** over their personal data such as the right to be informed what data is held about them, who has access to it, how it’s stored etc, and the right to request a copy.

- Personal data is anything that allows a **living** individual to be identified directly, or in combination with other information such as:
  - Name, address, phone number, email address
  - What they look like
  - What they do for a living and what they earn
  - What their relationship is with another person
  - What their hobbies are
  - What their opinions are and the opinions of others about them

- Applies to oral history interviewees but importantly to **any identifiable living individual** mentioned in an interview.
Legally ‘processing’ personal data: key principles for data controllers/processors

- Personal data should be ‘processed’ ie collected/stored/accessed: **lawfully, transparently** and **securely**, and for…

- **Legitimate purposes**: ‘performance of a task in the public interest’

- Data processed should be **relevant** and **accurate** and…

- Be retained for no longer than is necessary for the original purpose **but**: personal data may be stored for longer periods if it will be processed solely for ‘archiving purposes in the public interest’, scientific or historical research purposes or statistical purposes
GDPR exemptions for oral history+archives

• Every instance of personal data processing must be justified with a ‘legal basis for processing’ and you can only assert one legal basis for each type of processing.

• Any organisation or individual which collects and archives personal data ‘for purposes having public value beyond the immediate interests of the organisation itself’ as part of its ‘legitimate interests’ or for the ‘performance of a task in the public interest’, can use this as the first legal basis for processing personal data.

• Then if processing personal data in a certain way (such as recording and preserving oral histories) is part of your public and legitimate purpose you can then use the ‘Archiving in the Public Interest’ exemption as the legal basis and purpose for processing sensitive personal data (‘special category data’).

• Warning: If your legal basis is ‘consent’ and this is subsequently withdrawn you cannot substitute another legal basis to continue processing that data. So BL/OHS advises against using ‘consent’.
‘Archiving in the public interest’ exemption

• Where data processing is for the purpose of ‘archiving in the public interest’ then the processing is exempt from:
  – A data subject’s **right to access** and confirmation of processing (*eg* you do not have to search your entire archive in response to a Subject Access Request for ‘everything you hold about me’)
  – Their **right to rectification** (*eg* you do not have to ‘correct’ a historical record, although you may wish to add commentary in the event of disputed material)
  – The need to **delete data after its original purpose has expired** (*ie* you can retain personal data indefinitely to maintain the historic record).
  – Their **right to restrict processing or object to processing** (*eg* you do not have to suppress access to an archival recording upon request, unless the processing is causing ‘substantial damage or distress’ to one of the persons mentioned in it or related to it in some other way)
What is ‘special category data’? [part 1]

• Personal data about **identifiable living people** where its public release is **likely** to cause ‘substantial damage and distress’ to those individuals including:

  • **Religious/philosophical**
    – An interview where an individual talks about the religious views and worshipping practices of their family.

  • **Political**
    – A recording where an individual comments on the political views and party membership of a former colleague.

  • **Sexual/sexuality**
    – An interview where an individual talks about his sexuality and personal relationship history.
What is ‘special category data’? [part 2]

• Trade union activities
  – An interview where a former business owner complains about the activities of the trade union at his business and names union members

• Corporate or industry
  – An interview where an employee talks about the unpublicised financial difficulties of the private company they work for

• Illegal/criminal/bad behaviour/bullying/malpractice
  – An interview describes a former manager as a known bully, alleging that they were responsible for the departure of a number of employees

• Race or ethnicity related
  – An interview where an individual talks about the racial background of another individual without their knowledge
What is ‘special category data’? [part 3]

• **War/violence/Northern Irish troubles/colonial military activity**
  – *An interview where a number of named individuals are reported to have had an affiliation with the IRA*

• **Medical or health related**
  – *A recording where an individual talks about the mental health of a friend, the treatment they received and the medication they took*

• **Scurrilous content/gossip/rumours**
  – *An interview where an individual questions the paternity of a man based on a rumour that his mother had an extramarital affair*
Processing ‘special category data’

• The processing of Special Category Data (previously called ‘sensitive personal data’) for archival purposes is covered in Section 4(a) of Schedule 1 of the Data Protection Act 2018 – ‘necessary for archiving purposes… in the public interest’.

• **Note:** It is allowable to collect and even publish quite sensitive personal information as part of an oral history recording without further considering a legal basis, as long as such processing would not cause ‘substantial damage or distress’ to any identifiable living person.
What does ‘substantial damage and distress’ mean?

• Financial loss
• Physical harm
• ‘A level of upset or emotional or mental pain that goes beyond annoyance, irritation, strong dislike, or a feeling that the data’s release is morally abhorrent’ (ICO)

**Note**: it is no longer an adequate defence under GDPR to remove offending content from a website under only a ‘notice and takedown’ policy: steps must be taken before it is made public to assess it for sensitive content
Practical steps for compliance - Projects, interviewers and archivists (1): Documentation

• **Information leaflet about the project**: to be issued prior to the interview explaining objectives and data/privacy policies

• **Interview participation agreement**: a pre-interview form to seek agreement from interviewees that they understand the purpose of the project, including how and where their data will be stored and used, and that their data will be processed under the ‘archiving in the public interest’ exemption of GDPR

• **Interview recording agreement**: a post-interview form (unchanged) seeking copyright assignment and documenting any embargos, closures or other access restrictions (clearly time-coded and dated)
Practical steps for compliance (2): Personal data sensitivity review

- **Project staff awareness and training:** *Sensitive data about identifiable living people not already in the public domain which might cause substantial damage and distress*

- **Interviewer:**
  - Ensure pre-interview participation agreement completed
  - Flag-up possible sensitive content in interview to manager using ‘Interview sensitivity review form’ (spreadsheet)

- **Project Manager/Archivist:**
  - Reviews, listens, decides whether to redact/embargo
  - Documents all decisions in ‘Interview sensitivity review form’

- **Interviewer:** Reviews and finalises Recording Agreement; updates content documentation (summaries)
<table>
<thead>
<tr>
<th>Track</th>
<th>Timecodes</th>
<th>Details of potentially redactable sensitive content</th>
<th>Date of First Review</th>
<th>Author of First Review</th>
<th>Decisions Taken</th>
<th>Date of Second Review</th>
<th>Author of Second Review</th>
<th>Decisions Taken</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>00:43:36 - 00:46:51</td>
<td>Story about manager dismissing colleague who was caught stealing</td>
<td>05/12/2018</td>
<td>Camille Johnston</td>
<td>Colleague is named. Consider potential level of distress for third party</td>
<td>12/12/2018</td>
<td>Mary Stewart</td>
<td>Close for lifetime of named colleague</td>
</tr>
<tr>
<td>8</td>
<td>01:20:56 - 01:25:20</td>
<td>Story about disciplinary hearing: line manager and VB were attacked by individual.</td>
<td>05/12/2018</td>
<td>Camille Johnston</td>
<td>Mention of sensitive employment issue. Third party not identifiable. No action</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>00:56:10 - 00:58:31</td>
<td>Comments on colleague having an affair with Councillor</td>
<td>05/12/2018</td>
<td>Camille Johnston</td>
<td>Information in the public domain (newspaper article). No action</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Practical steps for compliance (3): Developing a retrospective sensitivity review process for access to legacy oral history

- Utilise interviewer knowledge of their own interviews and curatorial knowledge of legacy content and context
- Acknowledge oral history collections have varying risks
  - life story vs subject specific vs public events
- Word-search across data sets using lexicon of root-words to identify potential instances of sensitivity (ten areas defined by GDPR) for further review
- Document the full process of decision-making to show due diligence
- Recognise that there are limited resources for retrospective review
- Recognise the value of a risk-based assessment approach
Retrospective sensitivity review methodology: word-searching existing documentation

• Searching for words and roots* of words to pick up potential sensitive data:
  
  **Example:** Corporate or work-related sensitive data: corrupt*, disciplin*, dismiss*, fired, fraud, incompeten*, sack*, tribunal, unprofessional, fiddle

• Content summaries require different approach to transcripts:
  – Summary: “mentions colleague fired for fraud”
  – Transcript: “Susan was sacked as she had her hands in the till and fiddled the books”
Redaction and record keeping

• Audio file
  – Create a new WAV file, mute the section
  – Versioning system to distinguish master WAV file from edited/redacted file (at BL we use M0, M1, M2 within filename)

• Documentation (summary, transcript)
  – Create a new document, remove sections – with explanatory note – and date when section will open.
  – Versioning system needed for documentation

• Catalogue entry
  – Note the closure details clearly in access restrictions field

• Internal record keeping
  – Need to update all internal systems accordingly
Sensitivity checking workflow

1. Staff awareness
2. Identify and document possible sensitive data
3. Listen to identified sections
4. Discuss and sign off decisions
5. Mute contentious sections in access file

Document the process throughout
Find out more

• Oral History Society GDPR advice:  
  http://www.ohs.org.uk/advice/data-protection/

• Includes a more detailed document on the outline approach to sensitivity reviews:
Thank you!

oralhistory@bl.uk
www.bl.uk/oralhistory
www.twitter.com/BL_OralHistory