

3- Experiences

You're working on an autoethnography project exploring the intersection of personal experiences and mediated social life. This involves creative writing and using material from social media.

Questions and Key Points:

1. **Publicly Available Data:**
 - **Question:** Is it acceptable to use publicly available data from social media for research?
 - **Key Point:** Even if data is publicly available, it may not have been put there for research purposes. You likely need to submit an ethics application and request a waiver of consent.
2. **Ethics Application:**
 - **Question:** What is required in an ethics application when using social media data?
 - **Key Point:** You need to consider if the data was intended for research and include details such as the source of data, the type of information collected, and the terms of use of the platform.
3. **Scope of Consent:**
 - **Question:** How does the scope of consent affect the use of social media data?
 - **Key Point:** The National Statement requires considering the consent scope. If data was not consented for research, request a waiver of consent.
4. **Privacy Policies:**
 - **Question:** How do privacy policies of websites affect research data usage?
 - **Key Point:** Review and comply with the privacy policies of the platforms. Obtain and highlight the relevant sections if allowed under the privacy policy for research use.
5. **De-identification and Redaction:**
 - **Question:** Can information be de-identified or redacted for research purposes?
 - **Key Point:** De-identification might not be appropriate if the data remains re-identifiable.
6. **Ethics Approval Process:**
 - **Question:** How to proceed with an ethics approval for using social media data?
 - **Key Point:** Submit an application explaining the source and type of data, terms of use, and request a waiver of consent for impractical scenarios like scraping large datasets.
7. **Institutional Policies:**
 - **Question:** Are there examples of institutional policies on data use for research?
 - **Key Point:** Institutional policies may evolve, like Monash's policy allowing the use of student data for research post-2016 with ethics approval.
8. **Practical Considerations:**
 - **Question:** What are practical considerations when contacting individuals for consent?
 - **Key Point:** For large datasets, it's impractical to contact each individual; hence a waiver of consent is usually granted. For small datasets, the ethics committee may require you to contact individuals.
9. **AI and Data Scraping:**
 - **Question:** What are the ethical considerations around platforms like META scraping data for AI models?
 - **Key Point:** This falls under data protection and management handled by teams like eSolutions, not the ethics committee.
10. **Evolving Policies:**

- **Question:** Does Monash have a stance on protecting its own data from being used by platforms?
- **Key Point:** This is managed by teams focused on data protection and management, ensuring compliance with evolving policies.

Full Transcript

This is auto-ethnography, my favourite. I just love auto-ethnography, exploring the intersection of personal experiences and mediated social life. You're going to be doing a lot of creative writing, and a lot of your life work occurs online, using material from social media. Previously, before 2018, there was the idea that if it's publicly available, you could do whatever you want with it and use it for your research.

At Monash, we've always maintained, both pre- and post-2018, that while it's publicly available, it probably wasn't put there for research purposes for you to get a publication or a thesis out of it. So, you need to put in an ethics application. Most likely, if you're using human data, images, videos, and so on, it will need a waiver of consent. Because you're not going to go back and contact unknown people, and you're covered. Other institutions before 2018 had the same idea: publicly available, go do whatever you want with it.

In 2018, the National Statement was revised to include the use of data publicly available on social media or the web. The advice in the national statement is to look at the scope of consent when using data that is out there. Did people put it out there for research purposes? Did they consent for it to be used for research purposes? If the answer is no, then you need to request a waiver of consent from the committee in your APEX application.

Of course, there are different levels of information available out there. Some publicly public, and that's fine. But you need to be familiar with the terms of use of the platform you're using data from, and their privacy policies. I've been dealing with an application where they wanted to scrape data from a webpage. They wrote to the mediators and hadn't received a response, so they assumed they could scrape the data. I went to the terms and references, and one of the items stated that you cannot scrape data from any of their pages. I wish people would read that.

So, privacy policy on each of the pages and terms of use are very important to read. Do the work yourself because it makes committee members very upset when you haven't done the work and you've left it up to them to click on a link and do the investigation. You might get some very angry comments.

If you were looking at a situation where you're taking commentary or reviews from social media about a particular issue, for example, people reviewing Airbnb properties, and you wanted to use that information, even though it's in the public domain, it would require ethics approval. I'm wondering how you would then proceed with an ethics approval in that case because the person who provided the review is not going to be contactable after the fact to ask for their permission.

Redacting or de-identifying information is probably also not appropriate because it's still within our research context. So that would be an application for using data that's existing on the web. It's either identifiable or re-identifiable information because many users may not use their real names but have usernames, and it's pretty easy to backtrack and work out who that person is.

Let's assume its re-identifiable information. The application will be for the use of this data. The questions on the form will ask you about where the data is coming from, what information you'll be

collecting or scraping, and there are options on the form about how you would seek consent. Does the privacy policy of that website allow you to use data for research purposes? If yes, download the privacy policy and highlight where it covers the use of the information for research purposes.

For example, at Monash, pre-2016, our privacy policy did not allow the use of student emails for recruiting them for research purposes or using their data for research purposes. In late 2016, that changed. Our student privacy policy now allows the use of assignments, data, and email addresses for research purposes with ethics committee approval. If that's the case for the website, download and highlight that it's allowed under the privacy policy or terms of use, and the committee will be fine with that. You will still need to request a waiver of consent. If you're scraping a large number of comments, it's not practical to contact each person for permission. So, you will need a waiver of consent, which is very likely to be granted.

When you have only a handful of people and you request a waiver of consent, and it's clear you have their contact details, the committee may say, "Why can't you just write to them and ask for permission?" It's not practical for you to contact millions of records.

Regarding Meta's new privacy policy, which allows them to scrape images and data from their platforms for use in their AI models like Lama, this is more of a question for e-solutions. They handle AI data protection, data management, and data security. It's not something the ethics committee gets into, but we would flag relevant data information from the library and e-solutions.