

Interviews Going Open!

Naomi Truan, Sophie Granger, Jo Lychnara

Keywords: Personal information · Pseudonymization · Guidelines · Small-scale project · Qualitative data · Open science · FAIR principles · Ensuring ethical practices

Abstract:

The project *Interviews Going Open!* addresses the challenge of **sharing sensitive interview data** within the humanities and social sciences. Interviews are a valuable source of information, but concerns about **privacy** and **ethical** considerations often hinder their open accessibility. Another common challenge is how to make “thick description” (a property of qualitative fieldwork) suitable for metadata (which is often criticized as oversimplification in qualitative research).

Key points:

- Pseudonymization is **not only about names**, but also indirect identifiers such as a person’s education and location.
- The goal is to **protect** the interviewees’ identity without losing any (in our case, sociolinguistic) information.
- The process supports a consistent, systematic approach, especially with multiple researchers involved.
- The outputs map out how we navigated a **small-scale project** on a tight budget in just four months.

Approach:

- By hand, case by case, for data integrity
- Mix common tools (Word, Excel) with digital methods (Python, Regex)
- Focus on **preserving utility = enabling data reuse**
- We use pseudonymization methods to convey the context and subtleties that were provided by the removed information.
- **Transparency:** all edits are documented

Two outputs in Open Access:

Workflow Document

Phase	Step	Time per step (h)	Time per phase (h)
Phase 1: Explore the data Main working file: transcript_annotated.docx	Step 1: Set up a working diary	0,1	19
	Step 2: Familiarize yourself with the data	18	
	Step 3: Reflect	1	
Phase 2: Systematically annotate sensitive details Main working file: sensitive_data.xlsx	Step 1: Categorize	21	30
	Step 2: Decide on the strategy	7	
	Step 3: Reflect	1	
Phase 3: Develop and implement the pseudonymization scheme Main working file: pseudonymized_data.xlsx	Step 1: Develop a consistent methodology	7	30
	Step 2: Finalize the pseudonymization	23	
Phase 4: Reformat transcriptions and make final files Main working file: transcript_pseudonymized.docx	Step 1: Implement and automate	8	16
	Step 2: Check and finalize transcripts	2	
	Step 3: Pseudonymize metadata	6	
	Step 4: Publish	1	
Phase 5: Document, archive, publish Main working file: whole folder	Step 1: Document	1	5
	Step 2: Ask for informed consent to make the (pseudonymized) data open	2	
	Step 3: Archive	1	
	Step 4: Publish	1	
Total phases			100
Administrative tasks and meetings	General administrative tasks and meetings	30	30
Total project hours			130

Full title: Interviews Going Open! How to Pseudonymize Sensitive Interview Data: A Detailed Step-by-Step Guide (With Time Stamps)

When to use this document: Before starting transcription
This document is meant for researchers before the start of the project, to estimate the time and costs when applying for grants.

Keywords: WHEN & WHY
<https://shs.hal.science/halshs-04743263>

Pseudonymization Scheme

Full title: Interviews Going Open! Pseudonymization Strategies: Protecting Data While Preserving Utility

When to use this document: After data collection and transcription
Can be used to pseudonymize any corpus containing sensitive information, whether it is a data gathered by a student’s project or outputs from previous research.

Keywords: WHAT & HOW
<https://shs.hal.science/halshs-04842999>