Ethically sharing qualitative data

Sebastian Karcher, Qualitative Data Repository McGill University Data Anonymization Workshop Series October 17, 2023

> **QDR** The Qualitative Data Repository

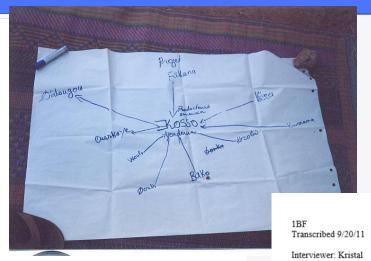
Shared Qualitative Data: Some Examples

Interviews, Focus Groups, and Ecological Data

README_Jones.txt (39.1 KB)

- ▶ Documentation
- • Qualitative Data
- **Farmer interview transcripts**
 - Burkina Faso
 - 🕨 🖬 Mali
 - Niger
- Village focus group seed maps
- • Village focus group transcripts
- Tabular data
 - Jones_Tabular_FarmerCharacteristics.csv (75.1 KB)
 - Jones_Tabular_RainfallByRegion.tab (15.0 KB)

Jones, Kristal. 2020. "Seed systems in West Africa". Qualitative Data Repository. https://doi.org/10.5064/F6URYY1I. QDR Main Collection.



Filename: Jones Tabular RainfallByRegion.tab

In Seed systems in West Africa (version 1.0), by Jones, Kristal

Download File **Close Preview**

Koumana		2010	2011	2012	Farmer ID	Region
	1	1104	920	1196	1	Dioila
	2	1104	920	1196	2	Dioila
	3	1104	920	1196	3	Dioila
•	4	1104	920	1196	4	Dioila
	5	1104	920	1196	5	Dioila
1BF	6	1104	920	1196	6	Dioila
Transcribed 9/20/11	7	1104	920	1196	7	Dioila
Interviewer: Kristal	8	1104	920	1196	8	Dioila
inciviewei. Klistai	9	1104	920	1196	9	Dioila
R: My French isn't very good.	10	1104	920	1196	10	Dioila
I: No, no, it's fine, it's fine.	11	1104	920	1196	11	Dioila
	12	1104	920	1196	12	Dioila
R: I didn't go to school. I only did three years.	13	1104	920	1196	13	Dioila
Background: That's why his French isn't very good.	14	1104	920	1196	14	Dioila
R: It's mixed up with Bomu. It's not good.	15	1104	920	1196	15	Dioila
	16	1012	828	1196	16	Siby
B:if it's not good, we can translate again.	17	1012	828	1196	17	Siby
R: That's already happened.	18	1012	828	1196	18	Siby
I: (laughs) Ok, and you are Bomu?	19	1012	828	1196	19	Sihv

R: Yes. Bobo.

I: Bobo, ok. And you are how old?

R: I was born in 1956.

I: 56. Ah, that's just like my mom! That's 54, no, 55. This year.

B: 55 years old.

I: And you speak Bomu, French?

Context for Ethnographic Research Videos and Online Sources



في ضيعة فايعة تأليف و سيناريو و حوار AForgoten Village Wedeen, Lisa. 2019. "Data for: Authoritarian apprehensions: Ideology, judgment, and mourning in Syria". Qualitative Data Repository. <u>https://doi.org/10.5064/F63776W4</u>

Listed URL	Full URL	Archived URL
nttps://tinyurl.com/y7rlby93	https://www.rt.com/news/39984	https://perma.cc/BC2C-SVEV
http://syriauntold.com/2018/06/o	http://syriauntold.com/2018/06/1	https://perma.cc/R8CV-6EL5
http://www.arabnews.com/news	http://www.arabnews.com/news	https://perma.cc/VVD6-27TH
https://tinyurl.com/yabcnmt8	https://view.publitas.com/the-int	https://perma.cc/7EFA-RDA9
nttps://tinyurl.com/lfnjcuw	https://www.nytimes.com/2013/	https://perma.cc/SGE3-AVWQ
https://tinyurl.com/y87dhbym	http://brown-moses.blogspot.co	https://perma.cc/LK6N-NC7L
https://tinyurl.com/ybbarngd	http://brown-moses.blogspot.co	https://perma.cc/TRF3-QFSQ
http://gawker.com/asma-al-assa	https://gawker.com/asma-al-ass	https://perma.cc/9MSA-VRCG
https://tinyurl.com/noh3fm9	https://www.nytimes.com/2013/	https://perma.cc/MEQ2-M7DZ
https://tinyurl.com/m9lh8q8	https://www.theguardian.com/w	https://perma.cc/764P-ZZXW
nttps://tinyurl.com/y8dxdswh	https://now.mmedia.me/lb/en/re	https://perma.cc/7R7M-X58T
https://tinyurl.com/nbrykrr	https://www.mintpressnews.com	https://perma.cc/BFW4-22VV
https://tinyurl.com/l835ygp	https://21stcenturywire.com/wp+	https://perma.cc/WWK7-ZSC4
https://tinyurl.com/ycyer6f4	https://www.gq-magazine.co.uk	https://perma.cc/V6DY-PEC5
https://tinyurl.com/y94gadfc	https://www.bellingcat.com/new	https://perma.cc/9HFX-EP44
https://tinyurl.com/ybaxukm2	https://www.bellingcat.com/new	https://perma.cc/KRP9-66R8
https://tinyurl.com/y7rld8ud	https://syriadirect.org/news/hts->	https://perma.cc/P9VN-BH8D

Mixed Method Data: Archival and Surveys

README_Hitt.txt (10.1 KB)

- <u>Hitt_Extracted Newspaper Excerpts.tab (</u>
 Book
 - Hitt_Ch1_Burson_Blackmun_1.pdf (12
 Hitt_Ch1_Burson_Blackmun_2.pdf (38
 Hitt_Ch1_Burson_Blackmun_3.pdf (3.8
 Hitt_Ch1_Burson_Blackmun_4.pdf (9.3
 Hitt_Ch2_Lockett_Blackmun.pdf (6.5 N
 Hitt_Ch4_Resweber_Jackson.pdf (6.1
 Hitt_Ch5_Tidewater_Jackson.pdf (8.6 I
 Hitt_Ch7 Watts Jackson.pdf (3.9 MB)
 - Replication
- Supplementary
 - The second seco
 - Adams
 - 🕨 🖀 API
 - Apodaca
 - Bailey
 - Baird
 - Bakke
 - Banco
 - Barr
 - 🕨 🖬 Beam
 - Bellotti
 - Books
 - Bowen
 - Braunfeld
 - Brock
 - Brown
 - Bullock

Anpreme Conrt of the United States Mashington, R. C. 20543

May 4, 1992

Re: No. 90-1056 - Burson v. Freeman

Dear Harry:

AUSTICE ANTONIN BEALIA

I withdrew my concurrence in this case (though I still considered that the preferable approach) in order to give you the fifth vote for a court. Since then (and indeed after the case was tentatively scheduled for release) you elevated footnote 11 to text, and altered it in a manner that I cannot accept. The new statement that the "First Amendment does not require States

to regulate for problems that do not ex intended, simply wrong. A state cannot dis for labor rallies alone, simply because that have historically produced violenc cannot go along with the new version's exclusive reliance upon the absence of or commercial solicitation has been use

Re: No. 90-1056, Burson v, Freeman

Dear Nino,

Hitt, Matthew. 2019. "Replication Data for: Inconsistency and Indecision in the United States Supreme Court". Qualitative Data Repository. https://doi.org/10.5064/F6W7QRSX

te cannot	distributionchannel	openingconsent	education	income	race	femal	e age_i	1 partyid	lean	strongdem	strongrep	ideology	polinteres	lawyer	sck1	sck2	sck3	scideo	pk1	pk2	pk3	pk4
d violenc ersion's	anonymous	1	4	4	1	1	28	3	2			5	3	6	0	0	1	3	secretary	50	1	8 years
	anonymous	1	2	2	7	1	69	3	2			4	2	6	1	0	0	3	don't know	don't know	0	4 yeasr
	anonymous	1	4	4	1	1	63	1			1	6	3	6		0	1	3	don't know	51		6 yrs
	anonymous	1	2	2	7	1	52	3	3			1	2	6		0		4	Don't know	48	1	4 yearz
	anonymous	1	3	2	1	1	46	2		1		3	3	6	0	0	1	4	Prime minister	100	1	8
sence of	anonymous	1	2	2	5	0	24	2		1		3	2	6	0	9	0	3	I don't know	200	0	2 years
been use	anonymous	1	2	3	3	1	46	3	3			4	1	6	1	9	0	2	sectarian	12		4 years
- Contraction and the second	anonymous	1	4	2	1	1	46	3	3			3	1	6		9		3	Don't know	Don't know		six years
	anonymous	1	3	3	1	0	46	3	3			7	2	6	0	1	0	4	leeword	50	1	2
	anonymous	1	3	2	1	1	39	3	3			4	1	6		9		2	Don't know	Don't know		Don't know
	anonymous	1	2	4	3	1	32	3	3			4	3	6	1	9	1	2	UK Prime Minister	435	0	2 years
	anonymous	1	4	3	1	1	24	2		2		4	1	6		0		3	?	435	1	2
	anonymous	1	2	5	8	1	53	1			2	7	1	6		9		3	don't know	don't know		don't know
	anonymous	1	2	3	1	0	26	2		1		1	2	6	1	9	0	3	British Prime Minister	435	1	6 Years
	anonymous	1	2	4	1	1	67	3	3			4	2	6		0	1	2	Don't know	Don't Know	1	I think 6 - not sur
	anonymous	1	4	5	1	1	27	1			1	7	3	6	1	1	0	2	International Relations.	125	0	8 years
	anonymous	1	4	1	1	0	32	3	3			4	1	6		9		3	I don't know.	I don't know.		I don't know.
	anonymous	1	2	5	1	0	69	2		2		4	2	6		0	1	2	unknown	100		6
	anonymous	1	3	2	1	1	61	1			2	3	3	6	1	1	1	3	prime minister - england	435	1	6 yr terms - unlto
	anonymous	1	3	3	1	0	26	1			2	4	3	6	1	0	0	2	conservative	535	1	6 years
	anonymous	1	2	2	1	1	36	3	3			4	1	6		9		2	Prime minister	435		6 years
	anonymous	1	4	5	1	0	34	1			1	3	2	6	1	1	1	2	dunno	300	1	10
	anonymous	1	2	1	7	1	56	2		1		6	2	6		0		1	Member of Parliament.	435	1	6 years
	anonymous	1	1	2	7	1	22	2		2		2	1	6		9		2	Don't know	Don't know		Don't know
	anonymous	1	4	5	5	1	28	3	3			1	1	6	1	1	0	1	Prime Minister of the United Kingdom	435	0	6 years
	anonymous	1	2	4	1	1	51	3	2			4	2	6	1	0		3	Member of Parliament.	435		6 years
		200		-																		

While I appreciate your having extended yourself to make a Court in <u>Burson</u>, I am unwilling to remove the two sentences you object to on p. 15 of the 5th draft. ("States adopt laws to address the problems that confront them. The First Amendment does not require States to regulate for problems that do not exist.") I believe that these sentences are necessary to respond to the dissent's contention that we have not "inquir[ed] into whether the content discrimination itself is related to the purported state interest" (p. 9 of 6th draft of JPS' dissent).

De-identified, Restricted Interviews & Documentation

- 1 CShdaimah Keisha Interview 1, Copy 2 12-4-12
- 2 INTERVIEWER:cause it would be sad for both of us. Okay. I'm gonna put this
- 3 just a little closer.
- 4 RESPONDENT: Okay.
- 5 INTERVIEWER: But you don't have to worry about speaking in it, it takes a...
- 6 RESPONDENT: Okay.
- 7 INTERVIEWER: So thank you again.
- 8 RESPONDENT: You're welcome.
- 9 INTERVIEWER: And I'm just gonna start off, I have just like a few demographic
- 10 questions, and then it'll be more like a conversation.
- 11 RESPONDENT: Okay.
- 12 INTERVIEWER: Okay. So how old are you?

13 RESPONDENT: I'm [30-40].

Shdaimah, Corey. 2020. "Problem-Solving Courts, Street Level Bureaucrats, and Clients as Policy Agents in a Prostitution Diversion Program". Qualitative Data Repository. <u>https://doi.org/10.5064/F6C8VUHP</u>

 Filename: Shdaimah_Guide_Interview-1.pdf

 Description: First Interview with Program Participants

 In Problem-Solving Courts, Street Level Bureaucrats, and Clients as Policy Agents in a Prostitution Diversion Program (version 1.0), by Shdaimah, Corey

 Download File
 Close Preview

Previous Next Page: 1/2

First Interview for Program Participants

A member of the research team will conduct the interview within first week of program acceptance. She will thank the participant for taking time to speak with her and review the letter of explanation. If the participant agrees to continue, the interviewer will stress that there are no right or wrong answers, and that she is interested in the participant's experiences with the Specialized Diversion Program/Project Dawn Court. The interviewer will remind the participant that she should not provide any identifying information in the course of the interview

Background

I am going to start with a few demographic questions:

- What is your age?
- What grade did you go to in school?
- · What do you consider to be your race or ethnicity?
- · What neighborhood or community do you come from in Baltimore?

Motivations

Researchers say that we don't really know enough about why women engage in prostitution. We think one of the reasons is that not many researchers talk to women. We hope it is okay to ask you these questions, but please feel free to refuse to answer them.

- · What was the major reason that you started to engage in prostitution? How old were you?
- · Are there any other reasons?
- Do you still engage in prostitution sometimes? If so, is this for the same reasons or other reasons?

Why Share Qualitative Data?

Why Share (Qualitative) Data?

- Be a good colleague: Allow others to scrutinize and build on your work
- Be a good teacher: Students benefit from working with "real" data
- Be a good citizen: Open science is more impactful science
- Get funded and published: Scientific funders and publishers increasingly expect data sharing and research transparency

Re-use in practice: Graduate Thesis

WestVirginiaUniversity

Graduate Theses, Dissertations, and Problem Reports

2022

Barriers and Opportunities to Exiting Prostitution: An Analysis of Prostitution Diversion Programs

Marina Binti Mohd Hamdan West Virginia University, mbm0031@mix.wvu.edu

https://researchrepository.wvu.edu/etd/11238

"This study is a critical program evaluation of two court-affiliated diversion programs with data collected by Corey Shdaimah from the University of Maryland School of Social Work. Shdaimah conducted a longitudinal study examining the perspectives of women exiting prostitution (...) through open-ended interviews with the participants as well as significant stakeholders (Shdaimah 2020).

I will be utilizing secondary data, her interview transcripts, of program participants from Specialized Prostitution Diversion program (SPD) in Baltimore City and Project Dawn Court program (PDC) in Philadelphia."

Re-use in practice: Developing & contrasting methodologies



The Qualitative Report

Volume 26 | Number 6

Special Section 17

6-22-2021

Introduction to Special Issue: Diverse Approaches to Qualitative Data Analysis for Applied Research

Jessica N. Lester Indiana University, jnlester@indiana.edu

Noah Goodman EDC's Center for Children & Technology

Michelle O'Reilly University of Leicester & Leicestershire Partnership NHS Trust, mjo14@leicester.ac.uk

https://doi.org/10.46743/2160-3715/2021.5015

"We suggest that the field can benefit from new and practical examples that allow researchers to compare different qualitative approaches, how they can be used in applied settings, and the unique lenses they bring. In order to achieve this goal, we have asked each author in this special issue to analyze one shared data set from a study exploring "postnatal care referral behavior by traditional birth attendants referral behavior by traditional birth attendants (TBAs) in Nigeria" (Chukwuma, Mbachu, Cohen, Bossert, et al., 2017). The shared data set, which was secured from Syracuse University's Qualitative Data Repository, consists of transcripts from three focus groups—one with hospital health care workers, another with traditional birth attendants, and a third with TBA delivery clients (Chukwuma, Mbachu, Cohen, McConnell, et al., 2017). Contributing authors—who work within in a range of disciplines were encouraged to focus less on the substantive findings of factors affecting TBAs' postnatal referral behavior, and instead aim to unearth the analytical contributions and insights that their approach might contribute to a similar data set."

Re-use in Practice: Teaching

From Alastair Tomlinson, Cardiff Metropolitan University:

"I teach a postgraduate research methods course for approximately 130 students per year, across a range of health-related degree subjects. We have used the QDR to enable students to develop their understanding of qualitative methods, and to demonstrate their qualitative data analysis skills. Students register with the QDR and access a specified research study. As part of their course assessment, students have to critique the qualitative design and methods used in the research study (relevance and appropriateness of design methodology to research question, evidence of rigour in sampling and data collection methods). They then access a subset of interview transcripts from the study and conduct a qualitative thematic analysis using the approach recommended by Braun & Clarke (2013). They produce a short research report summarising their evaluation and analysis, presenting the key themes that they have developed from their data analysis, and draw a brief conclusion. The QDR has been invaluable in enabling our students to work with real-world qualitative materials relevant to their subject area – and raising their awareness of the broader issues about making qualitative research data available to other researchers, and the appropriate access controls that need to be applied." controls that need to be applied."

Ethically Sharing Human Participant Qualitative Data

Ethics of Qualitative Research

- Qualitative research is often *relational:* The relationship between researcher & participant can be complex and multi-layered
- Often includes multiple interactions with varying levels of formality
- *Epistemological* commitments how we get to knowledge play an important role

Planning for Ethical Data Sharing

- Consider what and how to collect
- Consent for data sharing
- Manage securely
- De-identify as promised
- Control access as necessary
- Additional considerations depending on participant groups

Consider what and how to collect

- Don't collect identifying information if you don't need to
 - Example: Focus groups
- Careful with signed study information/informed consent
 - REBs/IRBs allow for oral consent when appropriate, but procedure needs to be documented
- Store identifying information separately from recordings/transcripts
- Plan for security, especially during fieldwork

Informed Consent

- Explicitly ask consent for data sharing
- Think about & ideally test appropriate language
- Don't assume a particular view by participants
- Opt-in (sometimes called 'tiered') consent can be helpful
- In some cases (e.g., First Nations/Indigenous groups), additional community consent is appropriate (more on that later)

Data Sharing in Informed Consent: Example

Alicia VandeVusse and Jennifer Mueller, Guttmacher Institute

Potential for Data Sharing: If you agree, the transcript of your interview may be shared with researchers at other organizations in the future. We will take out or change any information that could identify you before sharing. You can be in the study whether you agree to data sharing or not (see *Optional Consent* below).

Then include after the consent to participate: **Optional consent to data-sharing**:

Do you agree to allow a written copy of your interview to be shared with other researchers in the future?

- Yes
- No

In a qualitative study on abortion using this consent script, 92% of respondents opted into data sharing.

VandeVusse A, Mueller J, Karcher S. "Qualitative Data Sharing: Participant Understanding, Motivation, and Consent." *Qual Health Res*. 2022 Jan;32(1):182-191. doi: <u>10.1177/10497323211054058</u>

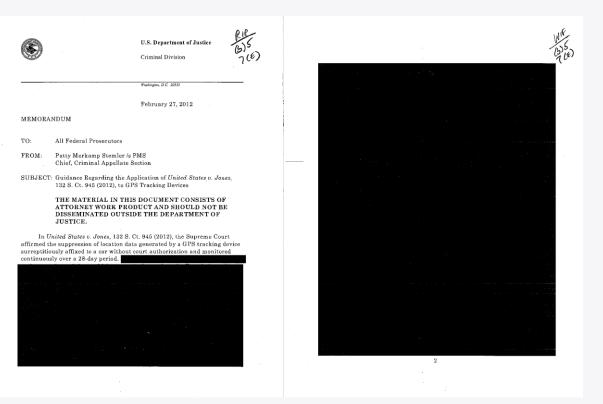
De-identifying Qualitative Research

De-identifying Qualitative Data: Between pointless and useless

4. Cohen's Illegal Campaign Contributions

On approximately June 16, 2015, Individual-1, for whom Cohen worked at the time, began an ultimately successful campaign for President of the United States. Cohen had no formal title

https://www.politico.com/f/?id=00000167-a496-df35-adef-fdf76fa30001



https://www.aclu.org/files/assets/doj_gps_tracking_memo1.pdf

Basic Guideline for De-identification

- Remove direct identifiers, or replace with pseudonyms often not essential research info
- Avoid blanking out; use pseudonyms or replacements (identify replacements)
- Consider broadening categories instead of redacting:
 - Born in 1937 → Born in [1930-1940]
 - Worked as a clerk in the zoning office \rightarrow Worked [for the city]
- Consider indirect identifiers in context
 - Removing one piece of information will affect how identifying another one is
 - Who has access to the data also affects re-identification risk

De-identification: Additional considerations

- Different disciplines/methods have different norms:
 - E.g., ethnography will almost always use pseudonyms and *may* accept "composite characters". Clearly marking redactions is less common (though debated)
 - In other fields, placeholders like "Interviewee 1" or "City 1" are sufficient
- For groups, keep a log to standardize de-identification
- Plan or apply editing at time of transcription or coding
- De-identification typically for persons, but can also be appropriate to mask locations or communities

De-identification exercise

- Go to: <u>https://managing-qualitative-data.org/files/de-identification-</u> <u>exercise.pdf</u>
- Take 15minutes to de-identify the text (assume/pretend it cannot be found online)
- We'll discuss together after

De-identification: A Solution

I was born in Philadelphia. My parents were both born and raised in Philadelphia. My father, **[Name]**, was Jewish and my mother, **[Name]**, was Irish Catholic. They both lived in South Philadelphia [...], and there was no chance that they would meet each other. Back in those days, and even when I was growing up, Philadelphia was a city of great ethnic divides, where the Italian, the Jewish, the Irish, the Polish, the black community, and—to the extent there was a Hispanic community—the Hispanic community each lived in their own neighborhood(s) with very little interaction.

They both went to the University of Pennsylvania, but didn't meet there. They met later on. They were both working in public assistance as social workers when they got married. The biggest thing was that back in those days an Irish Catholic was not very welcome in a Jewish family, and a Jew was not very welcome in an Irish Catholic family, so it was interesting growing up with these two ethnic backgrounds.

At Penn, my mother was president of her sorority and was a big person on campus. Interesting point, at that point the Daily Pennsylvanian, even though women had been there for a number of years, never had a woman's name in the newspaper. Even though they were students there, they were never mentioned. My mother went to [a girls high school] in South Philadelphia.

My father went to [a co-ed high school in South Philadelphia], and then went to Penn on a [sports] scholarship. [description of his role on the team] He thought he may have been one of the first Jews to play in the Ivy League. He played and started his first year, but he hurt his knee and lost his scholarship—which is what they did back then. His picture with his team [from the 1930s] is on the wall [at Penn]. He went back and earned a degree in fine arts at Penn. He then taught art in the city schools, and then returned to Penn and earned a master's degree in social work. He spent his career in social work and especially helping children. He finished his career as [working] for the City of Philadelphia. My mother worked in a number of social work jobs and later was a teacher in the Philadelphia City Schools. In [the 1940s] my parents had the first of my [...] wonderful sisters, [Martha], who we called [by a short version of that name].

Controlling Access

"As open as possible, as closed as necessary" (European Union)

- Default to open data (license "CCO" or "CC-BY")
- Public-use & restricted data
- Timed Embargo (in 1, 3, 10, 100 years)
- Access by application
 - Simple identity checks
 - Signed DUAs
 - Secondary IRB/REB
 - Depositor approved
- Access using enclaves (physical/virtual)

Additional considerations for research on:

- Groups with special status (prisoners, children, vulnerable groups)
- Indigenous groups
 - Legal requirements for research
 - Community consent
 - De-identification
 - CARE (Collective benefit; Authority to control; Responsibility; Ethics)/ Indigenous Data Sovereignty

Stay in Touch

• Email

- QDR: <u>qdr@syr.edu</u>
- Sebastian: <u>skarcher@syr.edu</u>
- Mastodon:
 - QDR: @qdrepository
 - Sebastian: @adam42smith
- Blog: <u>https://qdr.syr.edu/qdr-blog</u>