

Zertifikatskurs Forschungsdatenmanagement

FDM in den Sozialwissenschaften

10.11.2021

Seite 1


 This work is licensed under [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).


 Informationszentrum
 Lebenswissenschaften


 Landesinitiative für
 Forschungsdatenmanagement

ZBIW.

Technology
 Arts Sciences
 TH Köln

Wer sind wir?

Titel Vorname Nachname	Dr. Jan-Ocko Heuer
Tätig an folgender Einrichtung (seit)	FDZ Qualiservice (Universität Bremen), seit 2018
Im FDM tätig seit	2018
FDM-Aufgaben	Beratung und Schulung von Forschenden, Datenkuration, Entwicklung von FDM-Workflows, Entwicklung von FDM-Workshops/Curricula; Koordination des KonsortSWD-Measures TA.3-M.2 „FDM-Portfolio für die qualitative Sozialforschung“ in der NFDI
Meine Inhalte & Rolle im Zertifikatskurs	Dozent Modul 3 / FDM in den Sozialwissenschaften



10.11.2021

Seite 2

Wer sind wir?

Titel Vorname Nachname	Dr. Anja Perry
Tätig an folgender Einrichtung (seit)	GESIS – Leibniz-Institut für Sozialwissenschaften, seit 2012
Im FDM tätig seit	2016
FDM-Aufgaben	Beratung und Workshops zu FDM, Entwicklung von Domain-Data-Protokollen, Kompetenzentwicklung FDM in KonsortSWD, Leitung der DDI Training Group
Meine Inhalte & Rolle im Zertifikatskurs	FDM in den Sozialwissenschaften



Ablauf der Session



Was ist KonsortSWD?



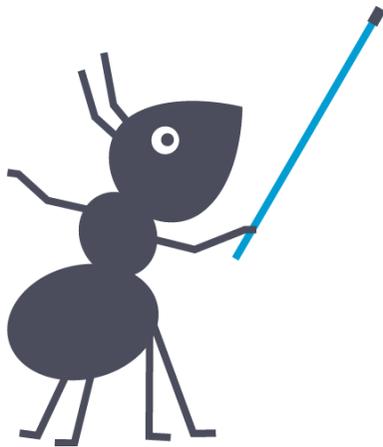
Datenschutz und Anonymisierung



Datendokumentation



Evtl. Fragen und Antwort



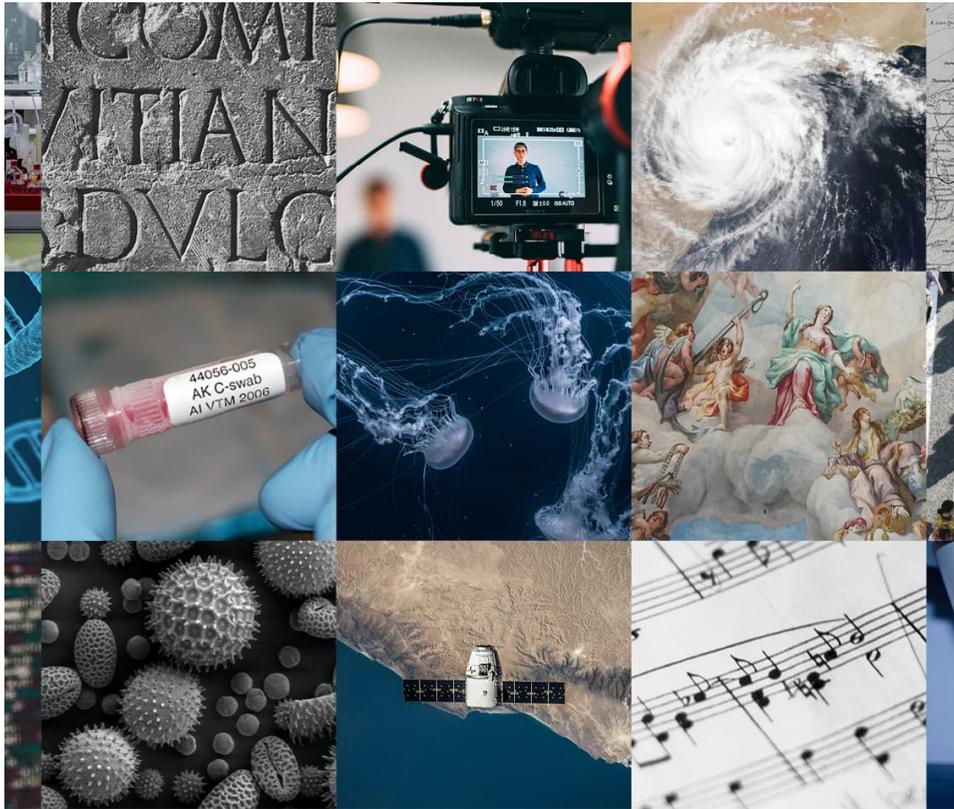
- Wir sind das NFDI-Konsortium mit der längsten Geschichte
- Forschungsdatenzentren (FDZ) seit 2001
- RatSWD seit 2004
- Unser Netzwerk galt als Referenz bei der Erfindung der NFDI

Die FDZ bei KonsortSWD





- Unterstützung der FDZ und stärkeres Zusammenwachsen
- Integriertes und zentrales Angebot für Forschende
- Neue Daten und Datentypen erschließen
- Auffindbarkeit verbessern
- Kompetenzen im Forschungsdatenmanagement aufbauen



Gibt es Fragen?

10.11.2021

Seite 8

Quantitative sozialwissenschaftliche Daten

	v1	v2	v3	v4	v5	v6	v7	v8	v9	v10	v11	v12	v13	v14	v15	v16
1	1.040e+14	4	26	104012	99	1	1973	2	1642	1040	1040	1	6	1	5	5
2	1.040e+14	1	19	104008	99	2	1943	1	99	99	99	1	1	6	4	5
3	1.040e+14	4	98	104003	99	1	1990	1	99	99	99	1	6	3	99	11
4	1.040e+14	2	98	104010	99	1	1983	2	1756	1040	1040	1	6	3	99	11
5	1.040e+14	1	18	104007	99	2	1927	1	99	99	99	1	1	6	4	1
6	1.040e+14	2	19	104007	99	1	1983	1	99	99	99	1	6	2	5	3
7	1.040e+14	4	15	104005	99	2	1970	1	99	99	99	1	1	2	4	6
8	1.040e+14	4	19	104006	99	1	1942	1	99	99	99	2	1	6	5	88
9	1.040e+14	4	15	104005	99	1	1965	2	1040	1276	1040	7	4	1	5	5
10	1.040e+14	7	15	104004	99	2	1955	1	99	99	99	1	1	2	5	5
11	1.040e+14	7	15	104003	99	2	7777	1	99	99	99	1	1	6	77	77
12	1.040e+14	4	18	104005	99	2	1938	1	99	99	99	1	3	6	5	5
13	1.040e+14	7	17	104005	99	1	1945	1	99	99	99	1	1	6	5	4
14	1.040e+14	4	18	104005	99	2	1949	2	1040	1380	1040	1	4	6	5	5
15	1.040e+14	2	15	104003	99	2	7777	1	99	99	99	1	1	2	4	5
16	1.040e+14	4	32	104007	99	1	1974	1	99	99	99	1	6	2	5	5
17	1.040e+14	4	98	104006	99	1	7777	1	99	99	99	1	6	3	99	11
18	1.040e+14	4	25	104010	99	2	1968	1	99	99	99	1	1	2	5	1
19	1.040e+14	4	40	104007	99	1	1967	1	99	99	99	1	1	2	6	2
20	1.040e+14	1	23	104004	99	1	1932	1	99	99	99	1	1	6	4	3
21	1.040e+14	2	18	104003	99	2	1965	1	99	99	99	1	6	2	5	5
22	1.040e+14	1	27	104011	99	1	1956	1	99	99	99	2	1	1	5	2
23	1.040e+14	4	18	104004	99	2	1923	1	99	99	99	1	3	6	3	3
24	1.040e+14	1	19	104006	99	2	1952	1	99	99	99	1	4	2	5	3
25	1.040e+14	3	16	104004	99	2	1947	2	1276	1040	1040	1	1	6	3	3

Quantitative sozialwissenschaftliche Daten

F003

ep04

⇒ Liste 3 vorlegen und bis Frage 4 liegen lassen!

Was glauben Sie, wie wird die wirtschaftliche Lage in Deutschland IN EINEM JAHR sein?
Benutzen Sie für Ihre Antworten bitte die Liste.

- Wesentlich besser als heute
- Etwas besser als heute
- Gleichbleibend
- Etwas schlechter als heute
- Wesentlich schlechter als heute
- Weiß nicht
- KA

F004

ep06

⇒ Liste 3 liegt vor!

Und wie wird Ihre eigene wirtschaftliche Lage IN EINEM JAHR sein?

- Wesentlich besser als heute
- Etwas besser als heute
- Gleichbleibend

```
Group
[id = BreakOffRouting | response condition = All | Hidden | layout = list ]

[BreakOffRouting]

routing
IF ('DISP_CI ≠ "1"') THEN Goto DC_CI_manual
ELSE
IF (('DISP_CI = "1"') AND ('DISP_BQ ≠ "1"')) THEN Goto DC_BQ_manual
ELSE
IF (((('DISP_CORE ≠ "1"') AND (('DISP_CORE ≠ "3"') AND ('DISP_CORE ≠ "4"')) AND ('COMPUTEREXPERIENCE = "1"')) AND (('CBA_Core_Stage1_Score = "NULL"') OR ('CBA_Core_Stage2_Score = "NULL"'))) THEN Goto DC_CORE_manual
ELSE
IF (((('DISP_CORE = "1"') AND ('CBA_Core_Stage1_Score > "3"')) AND ('CBA_Core_Stage2_Score > "2"')) AND ('DISP_CBA ≠ "1"')) THEN Goto DC_CBA_manual
ELSE
IF (((('COMPUTEREXPERIENCE = "2"') AND ('DISP_PP = "1"')) OR (((('COMPUTEREXPERIENCE = "1"') AND (('DISP_CORE = "3"') OR ('DISP_CORE = "4"')) OR ('CBA_Core_Stage1_Score ≤ "3"')) AND ('DISP_PP ≠ "1"')))) THEN Goto DC_PP_manual
ELSE
IF ((((((('COMPUTEREXPERIENCE = "2"') AND ('DISP_PP = "1"')) AND ('DISP_PRC ≠ "1"')) OR (((('COMPUTEREXPERIENCE = "1"') AND (('DISP_CORE = "3"') OR ('DISP_CORE = "4"')) AND ('DISP_PP = "1"')) AND ('DISP_PRC ≠ "1"')) OR (((('COMPUTEREXPERIENCE = "1"') AND ('DISP_CORE = "1"')) AND ('CBA_Core_Stage1_Score ≤ "3"')) AND ('DISP_PP = "1"')) AND ('DISP_PRC ≠ "1"')) OR (((('COMPUTEREXPERIENCE = "1"') AND ('CBA_Core_Stage1_Score > "3"')) AND ('CBA_Core_Stage2_Score ≤ "2"')) AND ('DISP_PRC ≠ "1"')) OR (((('COMPUTEREXPERIENCE = "1"') AND ('CBA_Start = "2"')) AND ('DISP_PP = "1"')) AND ('DISP_PRC ≠ "1"')))) THEN Goto DC_PRC_manual
ELSE Goto BreakOffError
```

Datenschutz und Anonymisierung

10.11.2021

Seite 11



ZBIW.

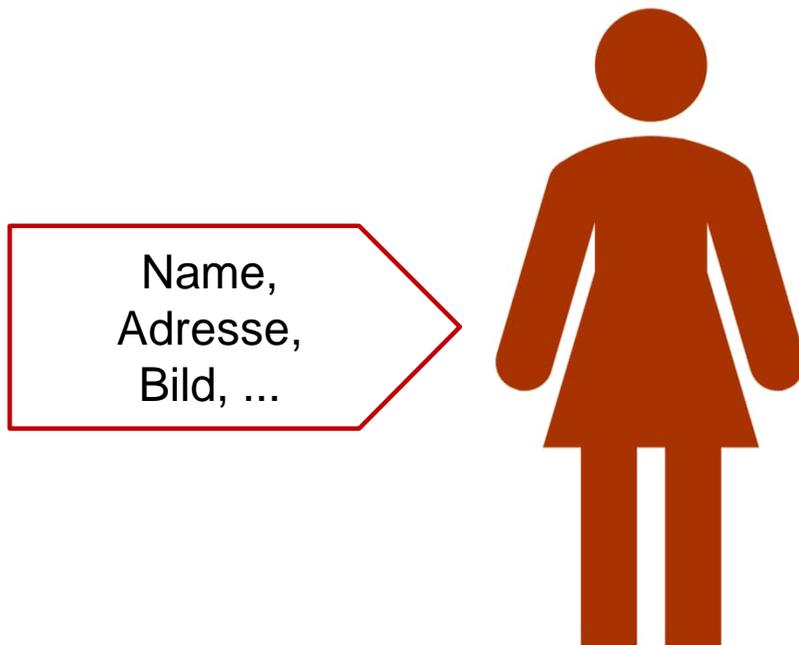
**Technology
Arts Sciences
TH Köln**

Datenschutz: Welche Informationen in den Fragen sind problematisch?

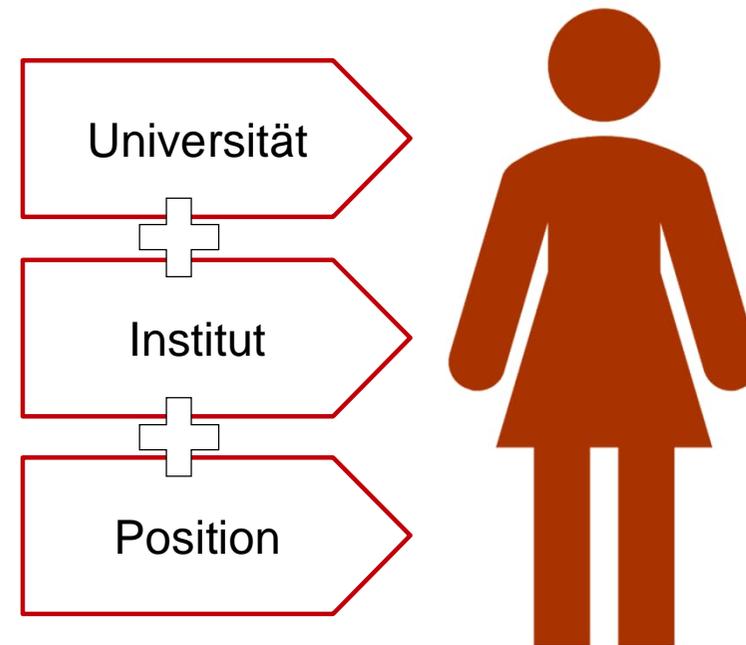
Nr.	Art der Information	Text	Antwort	Weiter mit
001	Frage	Geburtsdatum		002
		__/__/----		
002	Frage	Geschlecht		003
		1: Männlich		
		2: Weiblich		
		3: Divers		
		9997: Verweigert		
003	Frage	Familienstand		004
		1: Verheiratet		
		2: Eingetragene Lebenspartnerschaft		
		3: Geschieden/eingetragene Lebenspartnerschaft aufgehoben		
		4: Verwitwet/eingetragene(r) Partnerin verstorben		
		5: Ledig (NIE verheiratet oder in einer eingetragenen Lebenspartnerschaft)		
		9998: Weiß nicht		
		9999: Keine Angabe		
004	Frage	Leben in Ihrem Haushalt Kinder unter 16 Jahren, und wenn ja, wie viele?		005
		1: Ja, und zwar __ Kind(er) unter 16 Jahren		
		2: Nein		
		9999: Keine Angabe		
005	Frage	In welchem Land sind Sie selbst geboren?		ENDE
		1: Deutschland		
		2: Anderes Land, nämlich:		
		9998: Weiß nicht		
		9999: Keine Angabe		

Direkte und indirekte Identifikatoren

Direkt



Indirekt



Anonymisierung der erhobenen Daten

Kritische Variable	Risikierte Folge	Maßnahme
Geburtsdatum	Kann Identifizierung vereinfachen	Aggregation zu Jahr oder Monat und Jahr
Kinder im Haushalt	Hohe Kinderzahl kann Alleinstellungsmerkmal sein	Abschneidegrenze, z.B. "6 Kinder und mehr"
Herkunftsland	Seltene Einwanderungsland kann Alleinstellungsmerkmal sein	Seltene Länder gruppieren in Regionen
Familienstand	"eingetragene Lebenspartnerschaft" identifiziert Homosexualität	Variablenausprägung mit "verheiratet" zusammenlegen

Anonymisierungsstrategien

Direkte Identifikatoren immer getrennt von Umfragedaten aufbewahren oder löschen

Sensible Daten löschen, wenn nicht gebraucht, oder gar nicht erst erheben (Datenminimalismus)

Kombination von Merkmalen beachten

Abschneidegrenzen nach oben oder unten (z.B. bei Kindern im Haushalt, Einkommen)

Daten aggregieren (z.B. Bundesländer, Gemeindegrößenklassen, Berufsklassen)

Alle Änderungen dokumentieren!

Es kommt auf die Grundgesamtheit an!

Ausländer: Deutschland, Stichtag, Geschlecht/Altersjahre/
Familienstand, Ländergruppierungen/Staatsangehörigkeit

Ausländerstatistik Deutschland Ausländer (Anzahl)						
Stichtag Staatsangehörigkeit	Familienstand					
	ledig	verheiratet	verwitwet	geschieden	Lebenspartnerschaft	Lebenspartnerschaft aufgehoben
31.12.2020						
Insgesamt	4 961 395	4 767 230	216 105	612 670	10 365	1 645
davon:						
Ägypten	18 345	14 295	210	1 390	20	5
Äquatorialguinea	85	30	5	5	-	-
Äthiopien	12 495	5 320	215	760	40	-
Afghanistan	166 260	74 165	4 405	2 630	155	10
Albanien	32 870	31 310	360	1 740	30	5
Algerien	7 665	7 835	220	1 715	15	-
Andorra	15	5	-	5	-	-
Angola	3 390	1 290	80	335	5	-
Antigua und Barbuda	10	15	-	-	-	-
Argentinien	2 825	2 980	45	335	40	5

Quelle: Ausländerstatistik, GENESIS-Datenbank,
Statistisches Bundesamt

10.11.2021

Seite 16

Exkurs: k-Anonymität und I-Diversität

Definition

- Bestimmte Anzahl (k) von Personen, die die gleichen Merkmale besitzen, darf nicht unterschritten werden (k-Anonymität)
- Diese Gruppe darf gleichzeitig nicht die gleichen Ausprägungen haben (I-Diversität)

Vorteil

- Klare Regel, Entscheidungskriterium, wann ein Datensatz anonymisiert ist

Nachteil

- Informationsverlust

Angewendet bei Vollerhebungen und besonders sichtbaren Personen (z.B. Politiker*innen, Forschende, Personen des öffentlichen Lebens)

Haben Sie noch Fragen?

10.11.2021

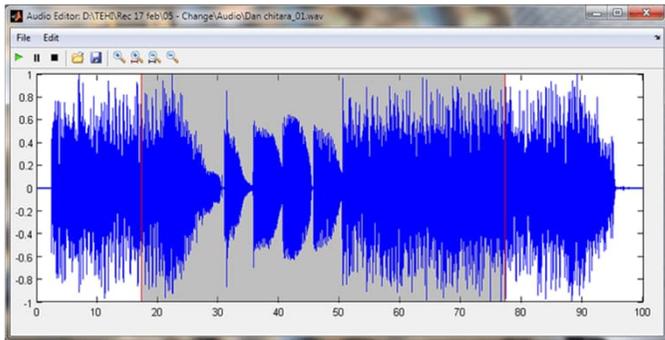
Seite 18



ZBIW.

**Technology
Arts Sciences
TH Köln**

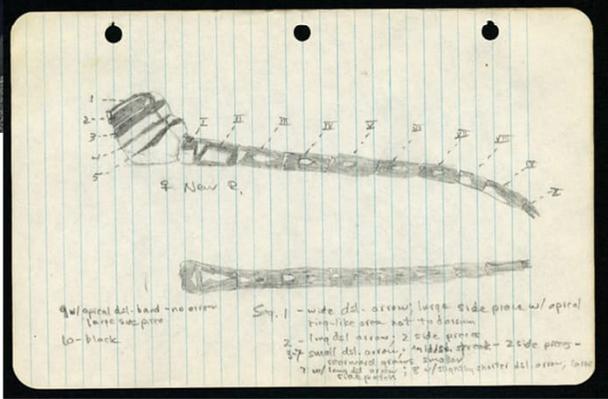
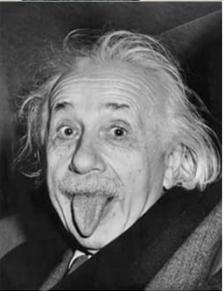
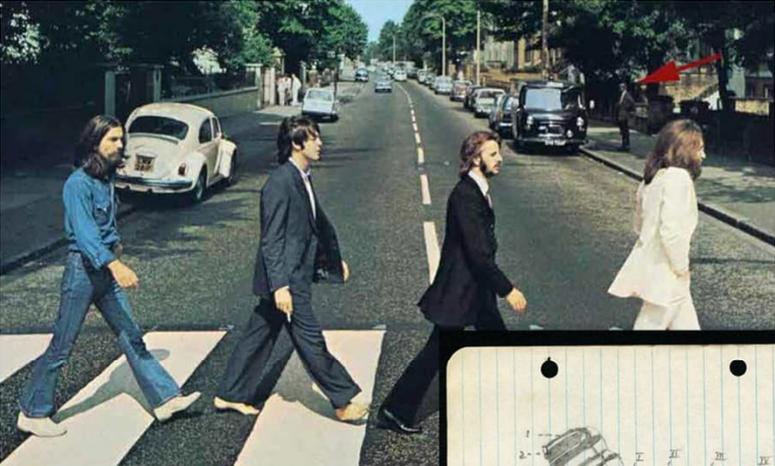
Beispiele für qualitative Forschungsdaten



Data	Codes
<p>(Q1) Moderator: What do you think about the apparel industry in Sri Lanka? Who are the customers, competitors, suppliers and influential parties and <i>what</i> is their influence on the business and management controls of the business?</p> <p>Finance Director: Global competition and GSP plus [Generalized System of Preferences] and the Trans-Pacific Trade Partnership have a direct impact on our operations, and on management controls. More than 60% of Sri Lankan apparel exports are to the United States and other Western countries. Our operations and controls have to focus on their demands in order to survive in the market. In the meantime, competition in the industry is increasing from emerging countries like Laos, Cambodia and Vietnam. So, our controls have always to focus on reducing costs and producing products at competitive prices. The economic environment of our country too is not supportive of the apparel industry because other industries such as tourism and hospitality are growing. Thus finding labour for the apparel industry will be difficult in about another five years because they expect</p>	<p>Global statutory rules</p> <p>Customer regulatory demands</p> <p>Influence of changing market conditions</p> <p>Facing changes in the local industry</p> <p>Influence of community</p>

Audio- oder Videoaufnahmen und Transkripte von Interviews, Fokusgruppen usw.

Beispiele für qualitative Forschungsdaten (III)



Bilder, Fotos, Zeichnungen, audiovisuelle Daten



Social Media Daten

Merkmale von qualitativer Sozialforschung & Daten

- Forschungsziele: oft Exploration, (dichte) Beschreibung, Theorieentwicklung
- Verstehen von Sinn – Analyse von Prozessen, Mechanismen, Konfigurationen etc.
- Deshalb: Offenheit des Forschungsprozesses (Fallauswahl, Methodenspezifikation)
- Resultat: viele, heterogene, sehr informationsreiche, relativ unstrukturierte Daten
- Besondere datenschutzrechtliche (und forschungsethische!) Herausforderungen
- Schwierigkeit I: Vertraulichkeit und Schutz von Forschungsteilnehmenden
- Schwierigkeit II: Informierte Einwilligung u.U. nicht möglich bzw. nicht schriftlich
- Schwierigkeit III: Anonymisierung von informationsreichen unstrukturierten Daten

Anonymisierungsübung im Selbstlernteil

Die Aufgabe lautete:

Markieren Sie im unter „Kursmaterialien“ verlinkten Auszug aus dem Transkript eines qualitativen (genauer: narrativ-biographischen) Interviews **jedes Wort, jede Phrase oder Textbestandteil, bei denen Sie eine Anonymisierung für erforderlich halten**, um eine Identifizierung der jeweiligen Person zu verhindern.

Überlegen Sie sich, wie bei der Anonymisierung vorgegangen werden könnte, um das Risiko einer Identifizierung von Personen zu verringern und zugleich Nachnutzungen der Daten zu ermöglichen.

Fragen / Diskussion zur Anonymisierungsübung

- Wie viele Stellen haben Sie gekennzeichnet?
- Bei welchen Stellen hatten Sie die größten Schwierigkeiten (und warum)?
- Wie könnte bei der Anonymisierung/Pseudonymisierung der markierten Stellen vorgegangen werden, um die personenbezogenen Daten der Personen zu schützen und zugleich (wissenschaftliche) Nachnutzungen dieser Forschungsdaten zu ermöglichen bzw. befördern?

Anonymisierung: zentrale Begriffe & Herausforderungen

- Keine Veränderung (ggf. andere Maßnahmen)
- Anonymisierung
- Pseudonymisierung
- direkte Identifikatoren
- indirekte Identifikatoren
- formal
- faktisch
- absolut
- Sensible Daten (z.B. „besondere Kategorien personenbezogener Daten“)
- Singuläre (oder seltene) Merkmale, Ereignisse etc.
- Kumulation von Informationen
- Verfügbarkeit von Kontext-Informationen (z.B. Publikationen, öffentlich zugängliche Quellen)

Das Konzept der „flexiblen Anonymisierung“

- Ziel: Balance schaffen zwischen Datenschutz und Ermöglichung von Nachnutzung
- Löschen, Schwärzen, „Verfälschen“ erschwert Forschung, führt zu falschen Resultaten
- Deshalb: Konzept der „flexiblen Anonymisierung“ (Pseudonymisierung) von Textdaten
 1. **Ersetzen (sensibler) personenbezogener Daten durch sozialwissenschaftlich relevante (abstrahierte) Informationen** (ggf. unterschiedliche Abstraktionsebenen, z.B. durch hierarchische Listen, Standardklassifikationen etc.)
 2. **Dokumentation der Ersetzungen** (= Pseudonymisierung), so dass Informationen auch wieder „geöffnet“ bzw. genauer spezifiziert werden können (und geschlossen)
 3. **Abhängig vom Forschungsinteresse der Nachnutzenden können bestimmte Informationen weniger pseudonymisiert/abstrahiert werden** („geöffnet“) und im Gegenzug andere Informationen stärker pseudonymisiert/abstrahiert werden, um Nachnutzungen zu ermöglichen und zugleich eine Identifikation zu vermeiden

Bsp.: Ersetzungskategorien bei flexibler Anonymisierung

	Ersetzungskategorie	Unterkategorien	Beispiel	
1	Ortsangaben	Land , Region, Bundesland , Gemeindeverband , Stadt , Stadtteil, Straßennamen, Gewässer, Gebirge Landstriche, Wohnort, Arbeitsort, Andere	Stadt Bremen	Stadt A, große Großstadt
2	Zeitangaben	Datum, Zeitspanne, Monatsangaben, Altersangaben	August 2002 – Oktober 2005	Sommer 2002 bis Herbst 2005
3	Organisationen & Institutionen	Firmen , Schulen , Ausbildungsstätten, Behörden/Verwaltung, Andere	Interviewdienst Horch	Unternehmen A - Unternehmen für sonstige freiberufliche, wissenschaftliche und technische Tätigkeiten (WZ M74)

Bsp.: Ersetzungskategorien bei flexibler Anonymisierung (II)

	Ersetzungskategorie	Unterkategorien	Beispiel	
4	Individuelle Besonderheiten	Krankheiten , Preise, sportliche Positionen, Unfall, Andere	Herzinfarkt	Ischämische Herzkrankheit (ICD10 I20-I25) oder Erkrankung des Kreislaufsystems (ICD10 I00-I99)
5	Personenbezogene Informationen	Anzahl der Kinder, Geschlecht, Name	Jutta	Name der Ehefrau
6	Berufliche und vergleichbare Tätigkeiten	Funktionen, Beruf, Tätigkeit , ehrenamtliches Engagement, Andere	Rettungs- sanitäter	Vertreter des Rettungsdienstpersonals (ISCO08 3258)
7	Ausbildung	Bildungsabschluss , Fach, Schulart , Weiterbildung, Ausbildungsform, andere	Altenpflege- ausbildung	Ausbildung zur Fachkraft für einen Assistenzberuf im Gesundheitswesen (ISCO3200)
8	Andere	Sprachen, Nationalität, Religion		

Tool für die (flexible) Anonymisierung von Text: QualiAnon

The screenshot displays the QualiAnon software interface. The top menu bar includes File, Edit, View, Anonymization, Analysis, and Help. The 'Analysis' menu is open, showing options like 'Create Replacement List ...' and 'Compare Codes ...'. The main window is divided into several sections:

- Categories:** A dropdown menu for 'Location' is open. Below it are input fields for 'Pseudonym', 'Description', 'Typ', 'Bundesland', and 'Gemeinde'. There are also buttons for 'C', 'S', 'Mark', 'Mark All', 'List Lookup', and 'Apply'.
- Replacements:** A table with columns: Originals, Category, Pseudonym, Replacement, and #. It lists entries for 'Katja Meier', 'Zwickau', and 'SPIEGEL ONLINE'.
- Project:** A list of files under 'Project: Mein Projekt', including 'Testinterview 1_Gruene.txt', 'Testinterview 2_Politik.txt', 'Testinterview 3_Wirtschaft.txt', 'Testinterview 4_Krankheiten.txt', and 'Testinterview 5_Russland.txt'.
- Search:** A search bar containing 'Meier' and 'Previous/Next' buttons.
- Text Content:** A list of numbered text snippets (5-13) with redaction boxes. For example, snippet (5) contains '[Person 1]' and '[Typ=Größere Mittelstadt]'. Snippet (7) contains 'erklärt'. Snippet (10) contains 'Wo kam das her?'. Snippet (12) contains '"Advent, Advent - ein Bulle brennt"'. Snippet (13) contains 'Man muss meine Punkvergangenheit zeitlich einordnen'.

Für Informationen und Zugang zum GitHub-Repository:

<https://www.qualiservice.org/de/helpdesk/webinar/tools.html>

Was gehört zu einer guten Dokumentation?

10.11.2021

Seite 30

Data Documentation

Go to
www.menti.com

Enter the code
6269 8233



Or use QR code

Worum geht es in diesen Daten?

	v1	v2	v3	v4	v5	v6	v7	v8	v9	v10	v11	v12	v13	v14	v15	v16
1	1.040e+14	4	26	104012	99	1	1973	2	1642	1040	1040	1	6	1	5	5
2	1.040e+14			104008	99	2	1943	1	99	99	99	1	1		4	5
3	1.040e+14			104003	99	1	1990	1	99	99	99	1				11
4	1.040e+14			104010	99	1	1983	2	1756	1040	1040	1				11
5	1.040e+14		18	104007	99	2	1927		99	99	99	1				1
6	1.040e+14	2	19	104007	99	1				99	99	1			5	3
7	1.040e+14	4	15	104005	99					99	99	1		2	4	6
8	1.040e+14	4	19	104005	99	1				99	99	2	1	6	5	88
9	1.040e+14	4				1			1040	1276	1040	7	4	1	5	5
10	1.040e+14	7				2	1959	1	99	99	99	1	1	2	5	5
11	1.040e+14	7				2	7777	1	99	99	99	1	1	6	77	77
12	1.040e+14	4				99	2	1938	1	99	99	99	1	3	6	5
13	1.040e+14	7	17	104005	99	1	1945	1	99	99	99	1	1	6	5	4
14	1.040e+14	4	18	104005	99	2	1949	2	1040	1380	1040			6	5	5
15	1.040e+14	2	15	104003	99	2	7777	1	99	99				2	4	5
16	1.040e+14	4	32	104007	99	1	1974	1	99	99				2	5	5
17	1.040e+14	4	98	104006	99	1	7777	1	99	99				3	99	11
18	1.040e+14	4	25	104010	99		1968	1	99	99	99		1	2	5	1
19	1.040e+14	4	40	104007	99		1967	1	99	99	99	1	1	2	6	2
20	1.040e+14	1	23	104004	99		1962	1	99	99	99	1	1	6	4	3
21	1.040e+14	2	18	104004	99		1965	1	99	99	99	1	6	2	5	5
22	1.040e+14	1	27	104011	99	1	1956	1	99	99	99	2	1	1	5	2
23	1.040e+14	4	18	104004	99	2	1923	1	99	99	99	1	3	6	3	3
24	1.040e+14	1	19	104006	99	2	1952	1	99	99	99	1	4	2	5	3
25	1.040e+14	3	16	104004	99	2	1947	2	1276	1040	1040	1	1	6	3	3

Wer wurde befragt?

Wie wurde erhoben?

Was bedeuten die Codes?

Wann wurde gefragt?

Wer hat gefragt?

???

Fukushima disaster: What happened at the nuclear plant?

🕒 10 March



Fukushima nuclear disaster

Inwieweit sind Sie für oder gegen den Gebrauch von Atomenergie in Deutschland?



<https://www.bbc.com/news/world-asia-56252695>

10.11.2021

Seite 33

Auffindbarkeit

Erwerbstätige in Zeiten von Corona (Februar/März 2021)

[Presse- und Informationsamt der Bundesregierung, Berlin](#)

GESIS Datenarchiv, Köln. ZA7697 Datenfile Version 1.0.0, <https://doi.org/10.4232/1.13802>

Abstract: Die Corona-Krise (COVID-19) und die damit einhergehenden Maßnahmen treffen einige Unternehmen und deren Beschäftigte hart. Vor diesem Hintergrund untersucht die Studie des Meinungsforschungsinstituts Kantar im Februar 2021 die persönliche Situation und die Arbeitsbedingungen von Erwerbstätigen in Deutschland in Zeiten von Corona. Die Analyse baut auf eine im Mai 2020 durchgeführte Befragung auf.

Primärforscher/ Wissenschaftlicher Beirat, Institution: - Presse- und Informationsamt der Bundesregierung, Berlin

Herausgeber: GESIS Data Archive

Studiennummer: ZA7697

DOI: 10.4232/1.13802

Publikationsjahr: 2021

Aktuelle Version: 1.0.0, 2021-10-08, <https://doi.org/10.4232/1.13802>

Erhebungszeitraum: 15.02.2021 - 01.03.2021

Untersuchungsgebiet: Deutschland (DE)

Grundgesamtheit: Deutschsprachige Bevölkerung ab 16 Jahren, die zum Zeitpunkt der Befragung in Voll- oder Teilzeit erwerbstätig ist

Untersuchungseinheit:  Individuum

Auswahlverfahren:  • Nicht-Wahrscheinlichkeitsauswahl: Quotenstichprobe; Quotenstichprobe (nach den Merkmalen Alter, Geschlecht, Bildung und Bundesland quotiert) aus Online-Access-Panel

Erhebungsmodus:  • Selbstausgefüllter Fragebogen: Webbasiert (CAWI);

Downloads

[Datensätze](#)

[Fragebogen](#)

[Andere Dokumente](#)

Aktionen

[Zitieren](#)

https://search.gesis.org/research_data/ZA7697

Was erwarten wir für die Datenveröffentlichung?

1. Fragebogen
2. Codebook
3. Methodenbericht
4. Studienbeschreibung

Und alles weitere, was für das Verständnis der Daten hilfreich ist.

Haben Sie noch Fragen?

10.11.2021

Seite 36



ZBIW.

**Technology
Arts Sciences
TH Köln**

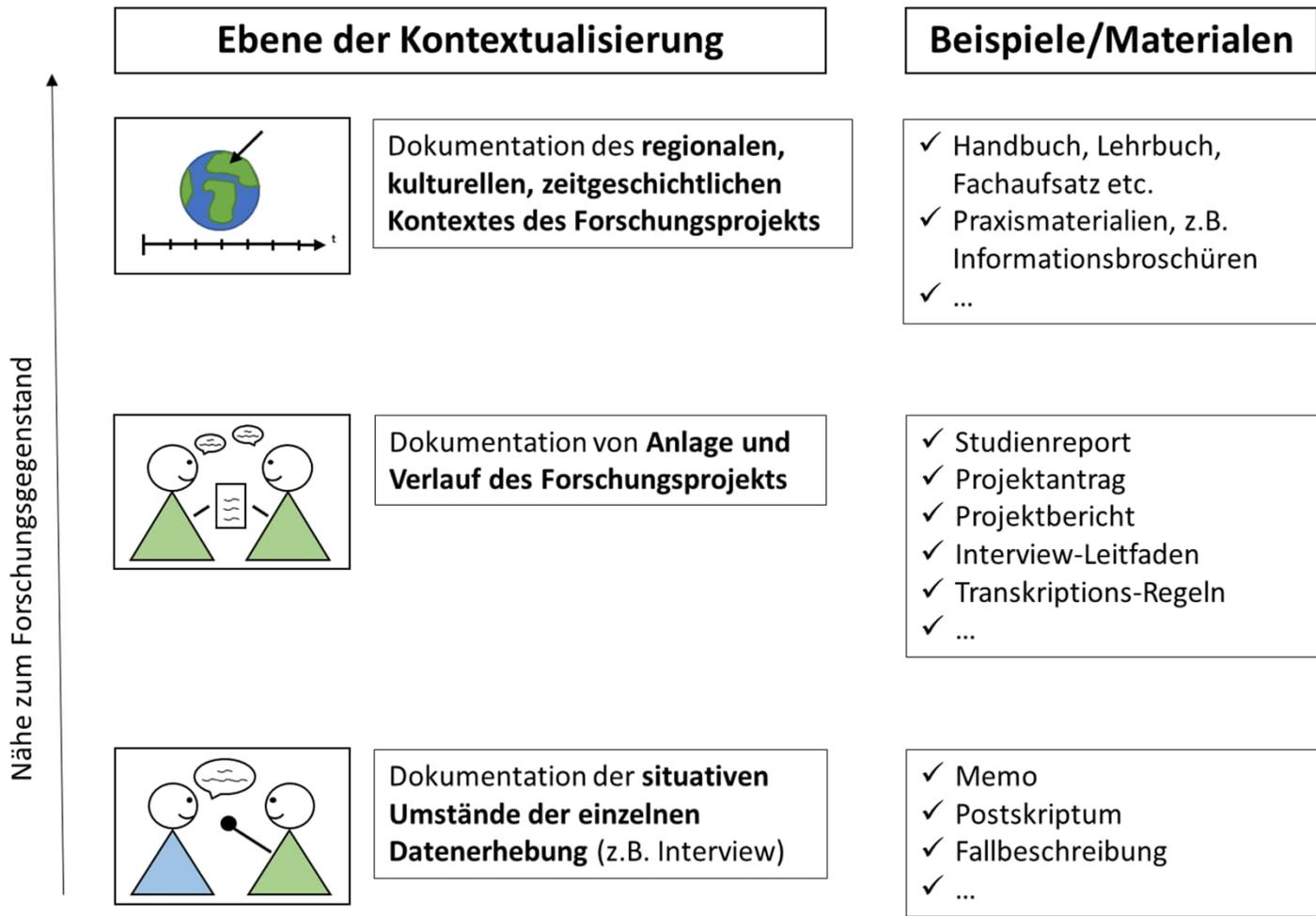
Merkmale von qualitativer Sozialforschung & Daten

- Qualitativ Forschende erheben ihre Forschungsdaten meist selbst („ins Feld gehen“)
- „Data Sharing“ ist – trotz (wegen?) des Informationsreichtums – noch die Ausnahme
- Rekapitulation: offener (und rekursiver) Forschungsprozess erfordert Modifikationen
- Rekapitulation: viele, heterogene, sehr informationsreiche, unstrukturierte Daten
- Deshalb: Besondere Herausforderungen auch bei der Dokumentation der Daten
- Weniger Standardisierung bei Methoden, Erhebungssituationen usw.
- Transparenz/Nachvollziehbarkeit (nicht Replikation) – und neue/andere Fragen!
- Generell: Dokumentation des Forschungsprozesses, der Datenerhebung und der Daten(aufbereitung) ist von zentraler Bedeutung!

Dokumentation qualitativer Daten: „Kontextualisierung“

- Dokumentation von qualitativer Forschung: mehr als „nur“ Metadaten
- Begriff für die Übermittlung relevanten Kontextwissens von Datenproduzenten (z.B. Primärforschung) zu Datennutzenden: KONTEXTUALISIERUNG
- Qualitative Forschung zielt auf sinnhaftes Verstehen von Handlungen bzw. Äußerungen in ihrem Bedeutungszusammenhang („Kontext“)
- Kontexte können unterschiedlich gut zu dokumentieren sein (z.B. informelles und intuitives Wissen, körperliches Erleben)
- Es gibt nicht *einen* Kontext, sondern vielfältige (z.B. institutionelle, inhaltliche, methodische und situative) Kontexte

Ebenen der Kontextualisierung (mit Beispielmaterialien)



Empfehlungen zur Kontextualisierung qualitativer Daten

- Es gilt, sich in die Situation von nicht am Forschungsprojekt beteiligten Wissenschaftler_innen hineinzusetzen und zu überlegen, welche Informationen diese zum Verständnis der Forschungsdaten benötigen
- Es gilt, eine Balance zu finden zwischen möglichst geringem Aufwand für Primärforschende und möglichst hohem Nutzen für Sekundärnutzende
- Idealerweise fortlaufend (d.h. als Bestandteil des Forschungsprozesses) eine Sammlung von (möglichen) Kontextmaterialien anlegen
- Jeden Schritt des Forschungsprozesses dokumentieren während er durchgeführt wird (weniger Zeitaufwand und höhere Qualität)
- Ideal (nicht nur für das „Data Sharing“): Verfassen eines „Studienreports“

Der „Studienreport“ als zentrales Kontextmaterial

- Für Nachnutzende ist es besonders hilfreich, wenn wesentliche Informationen in einem Dokument vorliegen („Studienreport“, „Daten-Methoden-Bericht“ o.ä.)
- Beim Erstellen können Datengebende in der Regel auf viele bestehende Materialien zurückgreifen (Anträge, Berichte, Publikationen etc.)
- Für Datengebende: Publikation --- für Nutzende: „Data fit“ einschätzen
- Vorschlag für die Strukturierung des Studienreports (an Bedarf anzupassen):
 1. Informationen zur Studie und zum Forschungskontext
 2. Datengenerierung: Vorbereitung und Durchführung
 3. Aufbereitung, Auswertung, Nachnutzungspotentiale der Daten
 4. Weitere Kontextmaterialien zur Studie bzw. zu den Daten

Haben Sie noch Fragen?

10.11.2021

Seite 42



ZBIW.

**Technology
Arts Sciences
TH Köln**