



Nachnutzungspotenzial erhalten und Persönlichkeitsrechte schützen: Die Anonymisierung qualitativer Daten als eine Datenschutzstrategie

Dr. Kati Mozygemba

QualidataNet, FDZ Qualiservice, Universität Bremen

DOI: <https://zenodo.org/uploads/17416399>

QualidataNet - Central Point of Entry für die qualitative Forschung



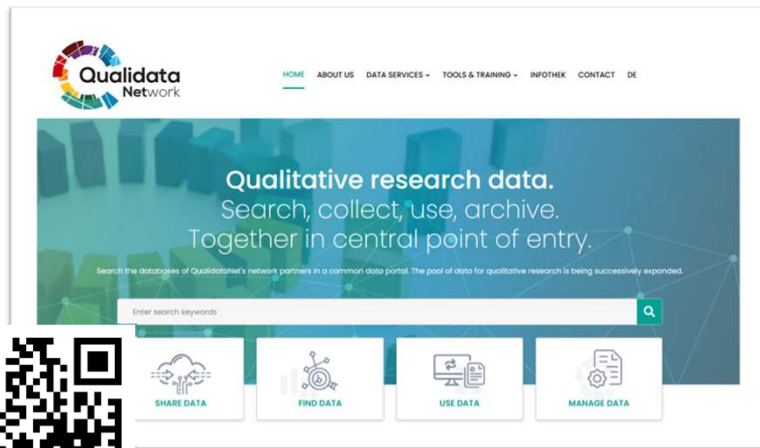
The screenshot shows the homepage of QualidataNet. At the top left is the logo, and to its right is a navigation menu with links: HOME, ÜBER UNS, DATEN SERVICES, TOOLS & TRAINING, INFOTHEK, NEWS & EVENTS, KONTAKT, and EN. The main content area has a blue background with a network diagram. The headline reads "Qualitative Forschungsdaten. Suchen, sammeln, nutzen, archivieren. Gemeinsam im Netz." Below this is a sub-headline: "Durchsuchen Sie die Datenbestände der Netzwerkpartner von QualidataNet. Der Bestand an Daten für die qualitative Forschung wird sukzessive erweitert." A search bar with the placeholder "Geben Sie einen Suchbegriff ein" and a magnifying glass icon is positioned below the text. At the bottom of the main area are four white boxes with icons and labels: "DATEN TEILEN" (cloud icon), "DATEN FINDEN" (magnifying glass icon), "DATEN NUTZEN" (document icon), and "DATEN MANAGEN" (gear icon).



QualidataNet – Netzwerk für qualitative Daten

Netzwerk von Forschungsdatenzentren, Archiven und Repositorien, die Forschungsdaten und -materialien der qualitativen Sozialforschung archivieren und für die wissenschaftliche Nachnutzung bereitstellen

Ziel: Bereitstellen eines central point of entry für qualitative Daten und (FDM-)Support



<https://www.qualidatanet.com/en/>

Services:

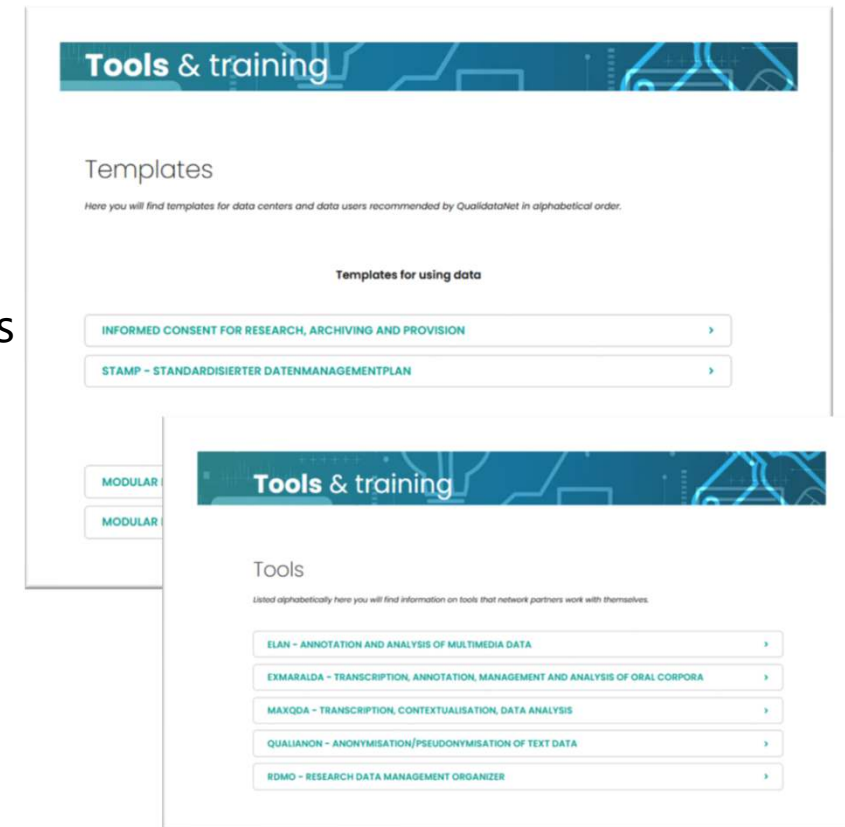
- Suchportal für qualitative Daten
 - Support in der Suche nach passenden Archivierungspartnern
 - Austausch und Integration themen- und datenspezifischer Expertise in gemeinsame Entwicklungen
- Z.B.: *das FDM-Portfolio für qualitative Daten*

Portfolio für das FDM qualitativer Daten

- FDM ist integraler Part der Forschungspraxis, weitreichender Support im Umgang mit Forschungsdaten
- Handreichungen und Vorlagen sind oft generisch >> fehlende Passung zum qualitativen Forschungsprozess >> Schwierigkeiten in der Anwendung
- Besonderheiten in der Datenaufbereitung mit Blick auf wissenschaftliche Nachnutzung
- **Forschungsfreundliches FDM muss die spezifischen Bedarfe qualitativer Forschung integrieren**
- **Archivierung und Nachnutzung sollte im gesamten Projektverlauf mitgedacht und umgesetzt werden**

Portfolio für das FDM qualitativer Daten

- Braucht eigene Perspektive und flexible Tools für das FDM qualitativer Daten
- Integrieren von Daten- und Projektspezifik
- Website: Übersicht über best practices, tools and templates
- Wir ergänzen gerne!
- Weiterführender Link/ detailliertere Präsentation:
<https://doi.org/10.5281/zenodo.15835923>
- Aktuell: Vorbereitung der Publikation einer Handreichung zum Data Sharing und FDM qualitativer Daten:
<https://doi.org/10.26092/elib/4334>



<https://www.qualidatanet.com/en/tools-training.html>

Zentral für das FDM qualitativer Daten ist ein Verständnis davon wie diese Daten charakterisiert sind

- Sensibilität der Daten und Themen
- Offenheit des Forschungsprozesses
- Kontextualität und (Verweisungs-)Dichte des Materials
- Zirkularität des Forschungsprozesses
- Diversität der Methoden und Datentypen
- Individualität und Projektspezifik
- Vertrauensvolle Integration in die Lebenswelt der Befragten
- Forscher identifizieren sich als Forschungsinstrument, Ko-Produktion von Wissen gemeinsam mit den Forschungsteilnehmenden

Was heißt das für die Anonymisierung qualitativer Materialien?

- Anonymisierung reduziert den Nachnutzungswert
- braucht studien- und datensatzbezogene Konzepte: Kein „one-size-fits all“ (Stam & Kleiner, 2020)
- Häufig wird von vorn begonnen
- Keine Standards wie z.B. k-Anonymität
- Dennoch häufig Orientierung an „Standard“-Ersetzungskategorien (Überanonymisierung und gering ausdifferenzierte Ersetzungsschemata)
- Häufig als Notwendigkeit angesehen oder in IC so formuliert
- Im Projekt Arbeit in der Regel mit pseudonymisierten Daten möglich (bis der Forschungszweck es erlaubt)
- Anonymisierung als iterativer Prozess

Konzepte und Techniken der Anonymisierung

- Anonymisierung ist eine Strategie in umfassenden Datenschutzkonzepten der FDZ, i.d.R. faktische Anonymisierung
- Aufbereitungsstrategien sollten auf verschiedenen Ebenen ansetzen
- Technisch-organisatorische Maßnahmen
- Die konkreten Verfahren sind abhängig u.a. von den in der Einwilligung zugesagten Datenschutzmaßnahmen sowie vom Forschungszweck.
- i.d.R. werden verschiedene Techniken miteinander kombiniert
- Wahl und Anwendung der Techniken orientiert sich daran, die Persönlichkeitsrechte der Betroffenen zu schützen und gleichzeitig von sozialwissenschaftlich relevante Information zu erhalten

Beispielfragen für die Erstellung eines Anonymisierungskonzeptes

- Wofür sollen die Daten anonymisiert werden – z. B. für die wissenschaftliche Nachnutzung, für die Zitation in einer wissenschaftlichen Veröffentlichung.
- Welche Identifikatoren sowie sensitiven Merkmale sind vorhanden? (Merkmale entsprechend Art. 9 DS-GVO (besondere Kategorien personenbezogener Daten) sowie ggf. forschungsethisch zu bedenkende Merkmale)
- Seltenheit - Gibt es seltene Ausprägungen, sodass nur wenige Merkmalsträger infrage kommen?
- Sind zusätzliche verknüpfbare Daten-/Informationsquellen öffentlich und stehen allen (etwa öffentliche Websites) oder spezifischen Gruppen zur Verfügung?
- Wie können sensible Stellen im Material so ersetzt werden, dass sozialwissenschaftlich relevante Information erhalten und „falsche Fährten“ verhindert werden?

Häufig als sensibel eingeschätzte Informationen

Ggf. zu anonymisieren:

- Charakteristika der Person
- Ortsangaben
- Zeitangaben
- Ausbildung
- Berufliche u.ä. Tätigkeiten
- Organisationen/Institutionen
- Individuelle Besonderheiten
- ...

Achtung: Mit der Ersetzung einer Information verändert sich der Informationsgehalt eines Textes und damit auch der Personenbezug von Informationen!

Original	Ersetzungsvorschläge
Ich bin Altenpflegerin.	<p>Ich bin [[Nennung des Berufs]].</p> <p>Ich bin [[Pflegerin/ Pflegekraft]]</p> <p>Ich bin [[Nennung des Berufs, Dienstleistungsberufe und Verkäufer, ISCO-08: Code 05:]]</p> <p>Ich bin [[Nennung des Berufs, Betreuungsberufe im Gesundheitswesen, ISCO08 Code 532:]].</p> <p>Ich bin [[Nennung des Berufs, Betreuungsberufe im Gesundheitswesen, Haus- und Familienpfleger, ISCO08 Code 5322:]]</p>

Aus: <https://doi.org/10.26092/elib/2525>

Techniken der Anonymisierung

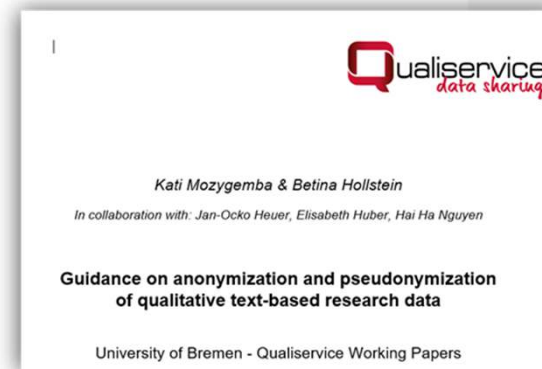
- die Abstraktion von Informationen
- die Aggregation von Informationen
- die relationale Ersetzung
- die kontextualisierende Ersetzung
- das Ersetzen durch „sprechende Pseudonyme“ (Butz, 2021)
- bedeutungsäquivalente sprechende Platzhalter (Meyermann & Porzelt, 2014, S. 7)
- das Paraphrasieren von ((berufs-)biografischen) Informationen
- Genderneutrale Ersetzungen (z. B. bei Gruppen mit geringen Grundgesamtheiten)
- das Löschen/Entfernen von Informationen.

Das Konzept der „Flexiblen Anonymisierung“

- Ersetzungen auf verschiedenen Abstraktionslevels
- Flexible Anonymisierung ist die Basis für das Anonymisierungstool QualiAnon
- Unterschiedlich differenzierte Ersetzungsschemata und unterschiedlich stark abstrahierte Versionen eines Textes können kreiert, erstellt und geteilt werden
- Möglichkeit, Ersetzungen zu öffnen oder zu schließen je nach Forschungsfrage oder Anwendungsszenario (z.B. onsite-Nutzung oder Lehre)

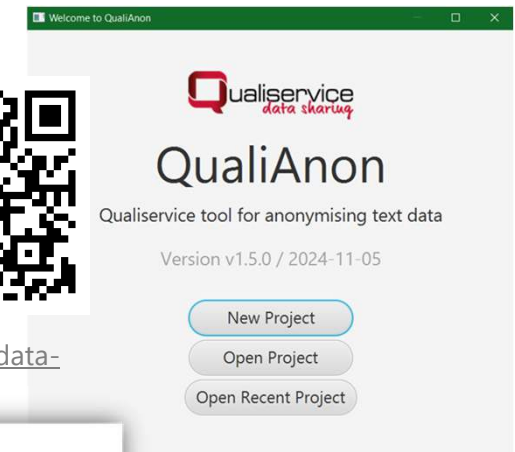


<https://github.com/pangaea-data-publisher/qualianon>



German Version: <https://doi.org/10.26092/elib/2525>

English Version : <https://doi.org/10.26092/elib/4029> (pre-registered)



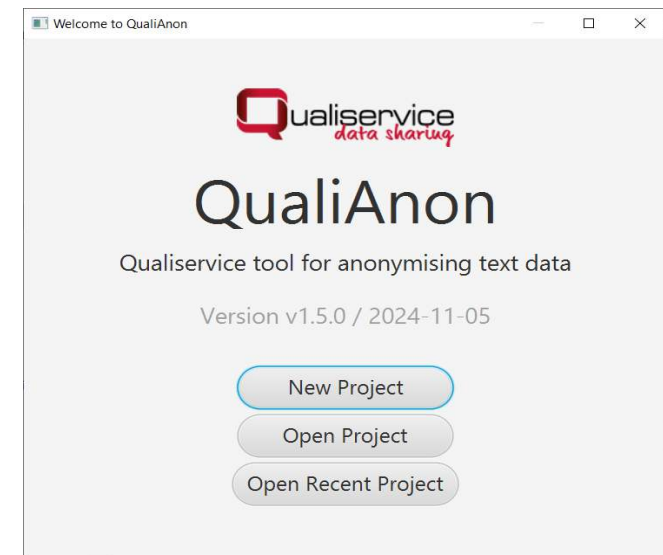
<https://www.qualiservice.org/en/the-helpdesk/tools.html>

Beispiele Ersetzung auf verschiedenen Abstraktionsebenen

Kategorie	Subkategorie	Beispiel	Replacement			
		Original	Level 1	Level 2	Level 3	Level 4
Orte	Land, Region, Bundesland, Stadt, Straße etc.	Berlin	Ort	Ort 1, Stadt in Deutschland	Ort 1, große Großstadt in Deutschland	Ort 1, große Großstadt in Deutschland mit mehr als 1 Mio. Einwohner
berufliche Tätigkeiten	Ausbildung, Berufe, berufliche Tätigkeiten	Altenpfleger	Beruf des Interviewten	Beruf des Interviewten, Dienstleistungsberufe und Verkäufer, ISCO-08: Code 05	Beruf des Interviewten, Betreuungsberufe, ISCO-08 Code 53	Beruf des Interviewten, Betreuungsberufe im Gesundheitswesen ISCO-08 Code 532
Individuelle Charakteristik	Krankheiten, Preise, Unfälle...	Schwerer Herzinfarkt	Krankheit	Krankheit 1	Schwere Erkrankung 1	Schwere Erkrankung des Kreislaufsystems

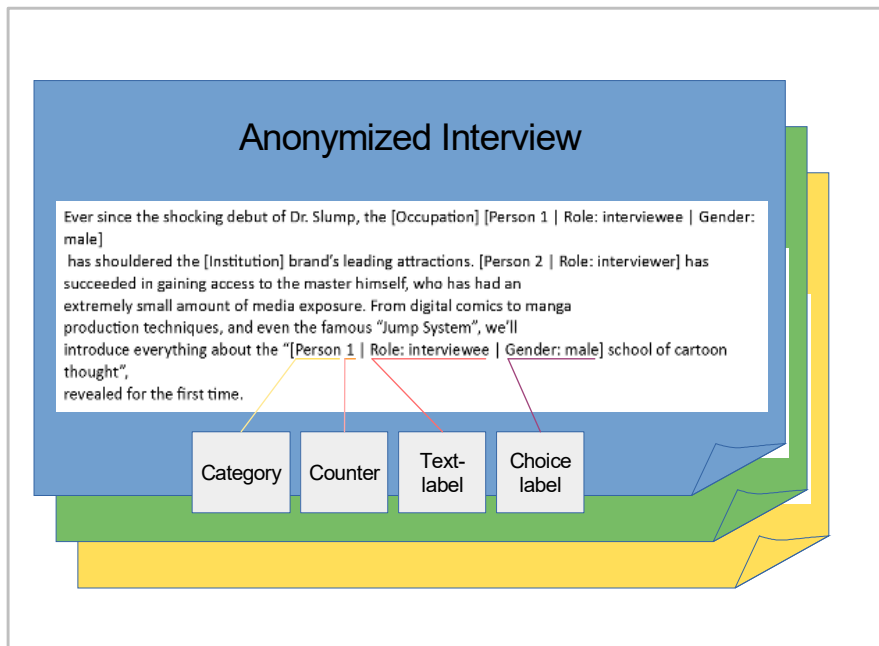
Tool für die Anonymisierung von Textdaten - QualiAnon

- Markieren und Ersetzen von sensiblen Informationen
- Halb-automatisch; Kontrolle durch Nutzende
- Dokumentation der Ersetzungen
- Nutzung von Listen/Klassifikationen für Ersetzung
- Beim Export: Auswahl des Abstraktionsgrads
- Separates Speichern der jeweiligen Dateien
- Anlegen, Teilen, Nachnutzen von Ersetzungslisten
- Open Source



Reference: Tom Nicolai, Kati Mozygemba, Susanne Kretzer, Betina Hollstein, Egor Gordeev (2025): QualiAnon - Qualiservice tool for anonymizing text data (version 1.5.0). Qualiservice. University of Bremen. Software available at: <https://github.com/pangaea-data-publisher/qualianon>

Beispiel für abstrahierten Text



https://docs.google.com/presentation/d/1GG_hlI0j7nXEFCsVKjJ44Tl1WQ0vOK7X/edit?slide=id.p1#slide=id.p1

- Für jede Entität, die QualiAnon ersetzt, kann aus einem bearbeitbaren System von Ersetzungskategorien, eine Kategorie gewählt (z.B. Person, Beruf, Ort, etc.) werden.
- In eckigen Klammern findet sich die Information, die zur Ersetzung verwendet wird:

hier: „[Person 1| Role: Interviewee| Gender: male]“.

- Nummerierung von Ersetzungskategorien

Blackened interviews
 Ever since the shocking debut of Dr. Slump, the [M:547fae83-53e1-455a-868a-38032e72c157:] [M:1d54e5a8-3af5-457a-9d36-37daadd99d55:] has shouldered the [M:2ca93753-0d37-4b17-e59c-e47f8065b3cb:] brand's leading attractions.

*- One UUID for each occurrence -
 Every marked area has its own unique ID*

Marker file for each blackened interview

Marker ID	Replacement ID	Original
[M:547fae83-53e1-455a-868a-38032e72c157:]	[M:05325-472e-4a25-a00c-71e2f54050c:]	caricaturist
[M:1d54e5a8-3af5-457a-9d36-37daadd99d55:]	[M:2ecb601e-c6bb-4afb-a588-968f99057e0:]	Akira Toriyama
[M:2ca93753-0d37-4b17-e59c-e47f8065b3cb:]	[M:a333d459-c1a8-4fca-bd18-e68cc66843ca:]	Jump

Connects the unique markers to common label data

One replacements file for whole project

ID	Role	Gender
[M:32b0e5ab-30a5-13ed-b438-fce12fa2636:]	interviewer	
[M:2ecb601e-c6bb-4afb-a588-968f99057e0:]	interviewee	male

*- One UUID for each entity -
 Label data is used consistently throughout project interviews*

Labels are defined in category schemes

- QualiAnon arbeitet mit einem System von **Universally Unique Identifiers (UUIDs)**
- Das anonymisierte Interview („blackened interview“) enthält die **Marker-ID**, die bei der Bearbeitung im Originaltext gesetzt wird
- Die **Replacement-ID** bezieht sich auf die jeweils gewählte Ersetzung
- In einem weiteren Dokument werden Marker-IDs und Replacement-IDs einer Studie sowie die Originalinformation gespeichert

https://docs.google.com/presentation/d/1GG_hlI0j7nXEFCsVKjJ44TI1WQ0vOK7X/edit?slide=id.p1#slide=id.p1

Demo Project - QualiAnon

File Edit Window Help

Documents

Replacements

Categories

Exports

interview-a

interview-b ×

Category / Originals	Labels	#
▼ Occupation		
cartoonist	Level 1: Professionals Level 2: Legal, social and cu... Level 3: Creative and perfor... Level 4: Visual artists	1
▼ Person		
Toriyama, Akira Toriyama, Akira Toriyama	Gender: Male Pseudo: Peter	33

1 **Akira Toriyama** Interview

2

3 Ever since the shocking debut of Dr. Slump, the cartoonist **Akira**

4 **Toriyama** has shouldered the Jump brand's leading attractions. Wired has

5 succeeded in gaining access to the master himself, who has had an

6 extremely small amount of media exposure. From digital comics to manga

7 production techniques, and even the famous "Jump System", we'll

8 introduce everything about the "**Toriyama** school of cartoon thought",

9 revealed for the first time.

10

11 Wired: It seems that the move to digital has begun, with even things

12 like comics being put on CD-ROM. Do you think that digital comics will

13 continue to increase in number from now on?

14

15 **Toriyama**: I think they'll probably keep increasing gradually. If we're

16 just talking about CD-ROMs, then ultimately it's almost entirely the

Log

Search: Toriyama

Markers: Akira Toriyama ×

- ▼ Person | Gender: Male, Pseudo: Peter
- ▼ interview-b
 - 1: Akira Toriyama Interview
 - 3: Ever since the shocking debut of Dr. Slump, the cartoonist Akira
 - 8: introduce everything about the "Toriyama school of cartoon thought",
 - 15: Toriyama: I think they'll probably keep increasing gradually. If we're
 - 27: Toriyama: Well, that in itself is obviously not a problem, as far as

12 like comics being put on CD-ROM. Do you think that digital comics

13 will

14 continue to increase in number from now on?

15 **Toriyama**: I think they'll probably keep increasing gradually. If we're

16 just talking about CD-ROMs, then ultimately it's almost entirely the

17 same work as reading comics [in print], since you're entirely free to

18 do things like go back [in the book]. CD-ROMs might even be better

19 in

20 that they don't get discarded and thrown away like paper comics.

Selected: 3
New replacement
Reassign
Remove

Saved document interview-b
Select Words



The screenshot shows the contact page of the Qualidata Network website. At the top left is the Qualidata Network logo. The navigation menu includes: HOME, ÜBER UNS, DATEN SERVICES, TOOLS & TRAINING, INFOTHEK, NEWS & EVENTS, KONTAKT (highlighted), and EN. Below the navigation is a teal banner with the word "Kontakt" in white. Underneath the banner is a heading "SENDE SIE UNS EINE NACHRICHT" flanked by horizontal lines. The main text reads: "Sie haben noch Fragen oder benötigen Hilfe? ... dann schreiben Sie uns gerne." Below this is a sub-heading "Unser Team freut sich auf Ihre Nachricht." followed by a paragraph: "Um Sie individuell unterstützen zu können, nehmen Sie gerne mit diesem Formular Kontakt zu uns auf oder schreiben direkt eine Nachricht an: info@qualidatanet.com". To the right of this paragraph is the address: "Qualidata Network, Universität Bremen, Mary-Somerville-Str. 7, D-28359 Bremen, Germany". At the bottom, there are three input fields: a wide one for "Betreff" (Subject), and two smaller ones for "Name" and "Vorname" (First Name).

www.qualidatanet.com

info@qualidatanet.com

