



Anonymisierung qualitativer Daten in der Kriminologie

1. Was sind qualitative Daten (in der Kriminologie)?

In den Sozialwissenschaften werden die in empirischen Forschungsvorhaben produzierten Daten in aller Regel in zwei verschiedene Bereiche abgegrenzt. So wird unterschieden zwischen quantitativen und qualitativen Daten. In Abgrenzung zu *quantitativen Daten* können unter dem Sammelbegriff *qualitative Daten* dabei all jene Daten zusammengefasst werden, die „im Rahmen von qualitativen Forschungsprozessen produziert und genutzt werden“.¹ Während im Grunde „alles“ Daten sein können bzw. als solche verstanden und analysiert werden können² und die Bandbreite der als qualitative Daten verstandenen Datenquellen demnach sehr groß ist, gibt es eine Reihe von mehr oder weniger standardmäßig in qualitativen empirischen Forschungsprojekten erhobenen Datentypen. Hierzu zählen vor allem Datenerhebungen im Rahmen von ethnografischen Studien oder Interviews bzw. Gruppendiskussionen. Die dabei verfassten Beobachtungsprotokolle sowie Audioaufnahmen und Transkripte von durchgeführten Interviews oder Gruppendiskussionen enthalten in der Regel eine Fülle von situativen Kontextinformationen und personenbezogenen Daten, die unter Umständen Rückschlüsse auf die Situation und/oder gar die betreffende Person ermöglichen. Während eine Re-Identifikation aus forschungsethischen wie auch datenschutzrechtlichen Gesichtspunkten grundsätzlich vermieden werden sollte, könnten eventuell mögliche Rückschlüsse auf eine bestimmte Person im Kontext kriminologischer Forschung schwerwiegende Folgen haben.

¹ Steinhardt et al (2020), S. 9.

² Das Diktum „all is data“ ist ein zentraler Kern der von Glaser und Strauss entwickelten *Grounded Theory Methodologie*, vgl. Glaser/Strauss (1998, passim).

In der deutschsprachigen Kriminologie stehen bei qualitativen empirischen Studien vor allem Audioaufzeichnungen bzw. Transkripte aus Interviews oder Gruppendiskussionen und in ethnografischen Feldforschungen entstandene Feldnotizen bzw. Beobachtungsprotokolle im Fokus. Da gerade im Kontext der Erforschung von abweichendem und strafbarem Verhalten eine Vielzahl an personenbezogenen und höchstpersönlichen Daten produziert und verarbeitet werden, kommt der gewissenhaften Anonymisierung dieser qualitativen Daten folglich eine hohe Bedeutung zu – insbesondere (aber nicht nur) im Kontext der Nachnutzung qualitativer Forschungsmaterialien.

2. Datenschutz als rechtliche Grundlage für die Anonymisierung

Empirische Forschungsarbeiten setzen voraus, dass mit den produzierten Daten verantwortungsvoll umgegangen wird. Während das für alle empirisch erzeugten Daten gleichermaßen gilt, erhält der Datenschutz bei Forschungsarbeiten zu abweichendem oder gar strafbarem Verhalten eine besondere Brisanz. Insbesondere bei qualitativer Forschung enthalten viele der generierten Daten solche Angaben, die nach § 3 Abs.1 BDSG zu den personenbezogenen Daten, also „Einzelangaben über persönliche oder sachliche Verhältnisse einer bestimmten oder bestimmbaren natürlichen Person“ zählen. Daneben sind die erhobenen Daten häufig auch „personenbeziehbar“, sodass diese „eine bestimmte Person zwar nicht eindeutig oder unmittelbar identifizieren, [...] aber erlauben, die Identität der Person mit Hilfe anderer Informationen festzustellen“.³ Aus diesem Grund zielt eine Anonymisierung der Forschungsdaten nach § 3 Abs. 6 BDSG darauf ab, personenbezogene Daten so zu verändern, dass die einzelnen Angaben „nicht mehr oder nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft einer bestimmten oder bestimmbaren natürlichen Person zugeordnet werden können“. Um diesen gesetzlichen wie forschungsethischen Vorgaben nachzukommen, soll die Anonymisierung der Daten im Forschungsprozess daher frühestmöglich erfolgen.

Aus forschungsethischen und datenschutzrechtlichen Gründen ist es erforderlich, dass die Forschungsteilnehmenden über den Sinn und Zweck der Datenerhebung sowie über den Umgang mit und die Verwendung der Daten umfassend aufgeklärt werden. Das Einverständnis zur Verarbeitung und Erhebung qualitativer Daten – beispielsweise im Rahmen eines qualitativen leitfadengestützten oder narrativen Interviews – ist daher noch vor Beginn der betreffenden Datenerhebung von den Studienteilnehmenden schriftlich einzuholen (Stichwort: „informed consent“). Sofern eine Nachnutzung der qualitativen Daten vorgesehen ist, muss dies entsprechend bereits im Zuge der Datenerhebung bedacht und den Studienteilnehmenden kommuniziert werden. Voraussetzung für die Nachnutzung ist hierbei, dass bei den Teilnehmenden im Zuge der Einwilligungserklärung (*informed consent*) auch deren Einverständnis hierzu eingeholt wurde.

³ Metschke/Wellbrock (2002), S. 19.

3. Vorgehen bei der Anonymisierung qualitativer Daten

Gerade weil qualitative Daten in der Regel viele persönliche Daten und auch dichte Kontextinformationen enthalten, zielt die Anonymisierung qualitativer Forschungsdaten darauf ab, die Identität der interviewten oder beobachteten Personen zu schützen. Dies ist ein zentraler Aspekt des Umgangs mit qualitativen Forschungsdaten. Zu beachten ist dabei jedoch, dass eine Anonymisierung immer mit einem Informationsverlust einhergeht. Gerade im Hinblick auf die Nachnutzbarkeit qualitativer Daten ist es besonders wichtig, Wege der Anonymisierung zu finden, die den interpretativen Gehalt möglichst erhalten und zugleich den Schutz der betroffenen Personen verlässlich gewährleisten. Denn, selbst bei starker Aggregation könnte die Re-Identifikation eines Mörders mit mehr als fünf Kindern aufgrund der besonderen Spezifika der schweren, dabei recht seltenen Straftat und einer häufigen Vaterschaft möglich sein – insbesondere, wenn wie hier zu vermuten auch noch medial über dieses Verbrechen und die besondere familiäre Situation des Täters berichtet würde.

Für die Anonymisierung qualitativer Daten kann grundsätzlich auf verschiedene Möglichkeiten zurückgegriffen werden. Im Folgenden werden die verschiedenen Anonymisierungsmodi vorgestellt und anhand von Beispielen illustriert.

3.1 Formale Anonymisierung

Bei der formalen Anonymisierung von qualitativen Daten werden personenbezogene Daten wie Namen, Orte oder Berufsbezeichnungen oder sonstige relevante Kontextinformationen, die eine Identifizierung der Person ermöglichen könnten, entfernt und mit einem Platzhalter gekennzeichnet.

z.B.

Günther Jauch arbeitet schon seit vielen Jahren erfolgreich als TV-Moderator von „Wer wird Millionär“ bei RTL.
[Name] arbeitet schon seit vielen Jahren erfolgreich als [Berufsbezeichnung] von [Name des genauen Arbeitsfelds/Bereich] bei [Name des Arbeitgebers].

Angela Merkel prägte das politische Geschehen im Rahmen ihrer Kanzlerschaft über 16 Jahre.
[Name] prägte das politische Geschehen im Rahmen der [politischen Position] über [viele Jahre, einen langen Zeitraum].

Durch die Löschung der Identifikationsmerkmale und die Ersetzung dieser durch einen Platzhalter können durch die bisweilen starke Abstraktion jedoch häufig für die Analyse wichtige Informationen verloren gehen.

3.2 Faktische Anonymisierung

Bei der faktischen Anonymisierung werden die personenidentifizierbaren Merkmale so verändert, dass die betreffende Person lediglich bei einem „völlig unverhältnismäßigen Aufwand re-identifiziert werden kann“.⁴

Daher werden hier Pseudonyme verwendet, nach denen die betreffenden Merkmale durch „Merkmale vergleichbaren Informations- und Sinngehalts“ ersetzt werden. Auf diese Weise können die pseudonymisierten Informationen weiterhin für die Analyse verwendet werden, sodass anders als bei der formalen Anonymisierung der Informationsgehalt erhalten bleibt. So wird bei der Auswahl der Pseudonyme beispielsweise darauf geachtet, dass Vor- und Nachnamen durch vergleichbare Namen mit ähnlicher Silbenanzahl und kultureller Zuordnung ausgewählt werden. Gleiches gilt für Pseudonyme bei Berufsbezeichnungen oder Herkunftsländern.

z.B.

Martina Mayer ist eine ausgezeichnete Friseurin.

Anita Müller ist eine ausgezeichnete Kosmetikerin.

Olaf Holm ist ein schwedischer Detektiv.

Sven Lund ist ein finnischer Polizist.

3.3 Aggregation von Informationen und Ersetzung durch ähnliche Bedeutungen

Daneben können zur Anonymisierung personenbezogener und personenbeziehbarer Daten auch die betreffenden Informationen durch Aggregationen zu bestimmten Klassen oder Kategorien zusammengefasst werden. Anstelle der starken Aggregation in große Zeiträume können insbesondere Jahres- oder Altersangaben in qualitativ empirischem Material auch durch einheitliche Veränderungen der exakten Angabe um z.B. zwei Jahre nach unten oder oben anonymisiert werden.

z.B.

Der heute 42-jährige Felix Neureuther ist erfolgreicher Skifahrer und kommt aus Garmisch-Patenkirchen.

Der heute zwischen 40 und 50 Jahre alte Mann ist erfolgreicher Sportler und kommt aus Süddeutschland.

Außerdem können die identifizierbaren Daten auch gemäß der jeweiligen Bedeutung oder Beziehung zu den Untersuchungspersonen anonymisiert werden.

⁴ Metschke/Wellbrock (2002), S. 21.

4. Dokumentation der Anonymisierungsvorkehrungen

Für eine stringente und einheitliche Anonymisierung sollten die unternommenen Anonymisierungsschritte in geeigneter Form dokumentiert werden. Dies ist insbesondere dann wichtig, wenn mehrere Personen für die Anonymisierung der Daten verantwortlich sind, um sicherzustellen, dass z.B. die Vergabe von Pseudonymen über die Daten hinweg einheitlich erfolgt. Das Anonymisierungsprotokoll darf nur unmittelbar beteiligten Forschenden zugänglich sein und muss getrennt von den anonymisierten Daten aufbewahrt werden.

5. “QualiAnon“ als hilfreiches Tool zur Anonymisierung qualitativer Daten

Praktische Unterstützung bei der Anonymisierung und Pseudonymisierung qualitativer Forschungsdaten bietet etwa das von Qualiservice zur kostenlosen Nutzung zur Verfügung gestellte Anonymisierungstool QualiAnon.⁵ Weitere Informationen zu QualiAnon sind unter: <https://www.qualiservice.org/de/helpdesk/webinar/tools.html> verfügbar.

Literaturverzeichnis

Glaser, B.G./Strauss, A.L. (1998): Grounded Theory. Strategien qualitativer Forschung, Bern: Verlag Hans Huber.

Kretzer, S./Mozygemba, K./Heuer, J-O./Huber, E. (2020): Erläuterungen zur Verwendung der von Qualiservice bereitgestellten Vorlagen für die informierte Einwilligung. Qualiservice Working Papers 2-2020, Bremen.
<http://dx.doi.org/10.26092/elib/192>.

Metschke, R./Wellbrock, R. (2002): Datenschutz in Wissenschaft und Forschung Verfügbar unter: https://userpage.fu-berlin.de/~makracht/mes/ws_2003/Unterlagen/Kursunterlagen/metadaten/mat_28_Datenschutz_Wissenschaft_Forschung.pdf [Letzter Aufruf: 7.7.2025]

Meyermann, A./Porzelt, M. (2014): Hinweise zur Anonymisierung qualitativer Daten. Frankfurt am Main: DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation. <https://doi.org/10.25656/01:21968>.

⁵ Nicolai et al (2021).

Weiterführende Literatur

- Mozygemba, K./Hollstein, B. (2023): Anonymisierung und Pseudonymisierung qualitativer textbasierter Forschungsdaten—eine Handreichung. In: Qualiservice Working Papers, 5. Bremen: Forschungsdatenzentrum Qualiservice.
<https://doi.org/10.26092/elib/2525>.
- Nicolai, T./Mozygemba, K./Kretzer, S./Hollstein, B. (2021): QualiAnon - Qualiservice Tool for Anonymizing Text Data (version 1.0.1). Qualiservice. University of Bremen. <https://github.com/pangaea-data-publisher/qualianon>.
- o.A. (2006): Anonymisierung von Daten in der qualitativen Forschung: Probleme und Empfehlungen. Erziehungswissenschaft, 17(32), S. 33-34.
- Steinhardt, I./Fischer, C./Heimstädt, M./Hirsbrunner, S.D./İkiz-Akinci, D./Kressin, L./.../Wünsche, H. (2020): *Das Öffnen und Teilen von Daten qualitativer Forschung: eine Handreichung*. (Weizenbaum Series, 6). Berlin: Weizenbaum, Institute for the Networked Society - The German Internet Institute.
<https://doi.org/10.34669/wi.ws/6>.
- Werner, C./Meyer, F./Bischof, S. (2023): Grundlagen, Strategien und Techniken der Anonymisierung von Transkripten in der qualitativen Forschung: eine praxisorientierte Einführung [78 Absätze]. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, 24(3), Art. 12, <https://doi.org/10.17169/fqs-24.3.4067>.

Stand: 7.7.2025