



Diese Use Cases sind als eine Hilfestellung für Forschende gedacht. Die Angaben im Text sind rechtlich nicht verbindlich. Bei Fragen zu Datenschutz, Urheberrecht und weiteren Rechtsaspekten wenden Sie sich bitte unbedingt an die zuständigen Stellen Ihrer Universität.

Die Use Cases sind in Zusammenarbeit der Open-Science-Teams der Universitätsbibliotheken Basel und Bern und der Datenschutzbeauftragten der Universität Basel nach der Durchführung eines Workshops zum Thema Datenschutz und Anonymisierung bei qualitativen Forschungsdaten entstanden. Beteiligte Personen: Silke Bellanger, Christina Besmer, Danielle Kaufmann, Iris Lindenmann, Jennifer Morger, Gero Schreier.

Der vorliegende Use Case steht unter einer [CC BY-SA 4.0](#)-Lizenz.

Zitieren als: Bellanger, S. [et al.]: Anonymisieren von Forschungsdaten.

Use Case: SNF, Open-Research-Data-Vorgaben, 30.10.2020, URL: https://researchdata.unibas.ch/fileadmin/user_upload/researchdata/Documents/UC_SNF-Vorgaben_20201030.pdf

SNF, Open-Research-Data-Vorgaben

Use Case

Noor Burhan forscht in einem Projekt zu den Auswirkungen der Finanzkrise von 2008 auf Schweizerische Industrieunternehmen. Noor Burhan arbeitet dafür mit Material aus Firmenarchiven und führt qualitative Interviews mit Mitarbeitenden der Firmen. Die Akten und Daten aus den Firmenarchiven enthalten Geschäftsgeheimnisse, sodass die Firmen eine Analyse erlauben, aber eine Veröffentlichung der Materialien ausschliessen. Die Interviewten haben nicht durchgängig eine Einwilligung zur Veröffentlichung ihrer Interviews gegeben, da diese unter anderem Fragen zu ihrer beruflichen Laufbahn enthalten und einige nicht möchten, dass diese Informationen bekannt werden. Aus den Transkripten werden im Projekt zwar alle direkten Identifikatoren entfernt, aber die Interviewten befürchten trotzdem, dass Personen aus ihrem Berufsumfeld sie erkennen könnten.

Das Projekt wird vom Schweizerischen Nationalfonds (SNF) finanziert. Gemäss der Open Research Data Policy des SNF wird erwartet, dass Forschungsdaten, die im Rahmen eines Projekts entstehen, spätestens zum Zeitpunkt der Veröffentlichung der darauf basierenden Publikation öffentlich zugänglich gemacht werden. Verlangt der SNF von Noor Burhan nun, alle Daten zu teilen?

Müssen Forschungsdaten von SNF-Projekten immer veröffentlicht werden?

Wissenschaftsförderer wie etwa der SNF verlangen, dass unterstützte Forschende alle Forschungsdaten, die ihren Publikationen zugrunde liegen, wenn immer möglich öffentlich zugänglich machen. Bei der Überlegung, ob und unter welchen Bedingungen Daten veröffentlicht werden können, sind jedoch rechtliche oder ethische Gründe, Vertraulichkeitsklauseln und Ähnliches zu berücksichtigen. Um den Schutz der betroffenen Personen zu gewährleisten, dürfen Personendaten, die nicht anonymisiert werden können und für die keine Einwilligung zum Veröffentlichen vorliegt, nicht veröffentlicht werden. Ähnlich verhält es sich zum Beispiel bei Geschäftsgeheimnissen

oder Materialien, die dem Urheberrecht unterliegen. Ausserdem können auch ethische Aspekte gegen eine Veröffentlichung von Forschungsdaten sprechen. Das heisst, in Fällen, in denen gesetzliche, ethische oder vertragliche Einschränkungen bezüglich einer Veröffentlichung von Daten oder Werken vorliegen, müssen diese respektiert und in den Datenmanagementplänen erläutert werden.

Was kann man tun, wenn Daten zum Beispiel aus rechtlichen oder ethischen Gründen nicht offen geteilt werden können?

Grundsätzlich ist Teilen nicht gleich Teilen. Es können zum Beispiel alle Daten unmittelbar auf einem Repository für alle öffentlich zugänglich gemacht werden oder unterschiedliche Bedingungen definiert werden, unter denen Daten eingesehen werden können. Manche Repositorien bieten in dieser Hinsicht spezialisierte Lösungen und kommen für das Teilen von heiklen Daten in Frage. Für die Sozialwissenschaften bietet zum Beispiel FORSbase Möglichkeiten an, Forschungsdaten mit unterschiedlichen Zugangsbeschränkungen zu veröffentlichen. Unabhängig von möglichen Zugangsbeschränkungen müssen für nicht-anonymisierte Personendaten aber immer entsprechende [Einwilligungen](#) vorliegen. Je nach Ausgestaltung der Einwilligungserklärung können dann zum Beispiel die Daten nur auf Anfrage und mit einer [Nutzungsvereinbarung](#) geteilt werden. Alternativ können, wenn die Einwilligungserklärung es erlaubt, die Daten verschlüsselt auf einem sicheren und regelmässig gewarteten universitären Server gespeichert werden und in einem öffentlichen Repository Metadaten zu den betreffenden Datensätzen hinterlegt werden.

Was sind Metadaten und welche Vorteile bieten sie, wenn Daten nicht geteilt werden können?

Metadaten sind Daten über Daten. Klassische Metadaten sind zum Beispiel Autor*in, Titel, Ort, Datum. Für empirische sozialwissenschaftliche oder historische Studien sind auch genauere Beschreibungen wie zum Beispiel Zeitpunkt, Ort und Länge der Interviews, Zahl der Studienteilnehmenden oder Erhebungs-/Befragungsmethoden möglich. Stich- und Schlagwörter erlauben eine Charakterisierung des Inhalts und eine thematische Einordnung, ohne dass sensible Details über Studienteilnehmende preisgegeben werden müssen.

Auch wenn Datensätze nicht öffentlich geteilt werden können, können Metadaten dabei helfen, die Datengrundlage, das Studiendesign, die Methodik und so weiter von Forschungen und Publikationen zu dokumentieren. So kann man grundlegende Informationen über Daten veröffentlichen und das Vorhandensein von Untersuchungen beziehungsweise Daten bekanntmachen, ohne die Daten selbst im Internet zu verbreiten. Dies erlaubt den Dialog mit anderen Fachpersonen, steigert die Sichtbarkeit der eigenen Forschung, kann zu mehr Zitationen führen und stellt einen weiteren Eintrag auf der eigenen Publikationsliste dar.

Spezialisierte Repositorien wie zum Beispiel FORSbase bieten Eingabemaschinen für eine detaillierte Beschreibung von Daten nach fachspezifischen Standards. In weniger spezialisierten Repositorien wie zum Beispiel [Zenodo](#) oder institutionellen Repositorien gibt es zu meist zusätzliche Freitextfelder, in denen man spezifische Metadaten unstrukturiert eingeben kann.

Was bedeutet das für diesen Fall?

Noor Burhan kann die Daten aus den Firmenarchiven sowie die Interviewtranskripte nicht veröffentlichen, da sie nicht vollständig anonymisiert werden können. Beim Erscheinen der abschliessenden Monographie trägt Noor Burhan jedoch im Repository FORSbase Metadaten zu der Studie ein. Der Eintrag erlaubt eine eingehende Beschreibung der Daten nach den Standards der empirischen Sozialforschung, an denen Noor Burhan sich während der Forschungsarbeit orientiert hat. Die Transkripte speichert Noor Burhan auf einem verschlüsselten Ordner auf dem universitären Server. Nach einiger Zeit meldet sich ein Forscher bei Noor Burhan, der auf den Metadateneintrag auf FORSbase gestossen ist und sich im Kontext eines eigenen Forschungsprojekts für die Originaldaten interessiert. Da einzelne Interviewpartner*innen ihre Einwilligung gegeben hatten, dass ihre Interviews auf Nachfrage zu wissenschaftlichen Zwecken in anderen Projekten verwendet werden können, kann der Forscher eine von Noor Burhan aufgesetzte Nutzungsvereinbarung unterschreiben und in der Folge diesen Teildatensatz auf einem sicheren Weg erhalten.