Personenbezogene Daten -
Definition, Speicherung und Minimierung

In der Forschung kann eine Verarbeitung (Art. 4 Abs. 1 DSGVO) von personenbezogenen Forschungsdaten stattfinden. Dies geschieht meist durch Beobachtung von Individuen oder bei der Abfrage von Informationen durch Umfragen. Diese Handreichung soll einen Überblick zu den wichtigsten Aspekten im Umgang mit personenbezogenen Daten geben. Dabei wird mit der Definition dieser Daten begonnen, erklärt was eine Einverständniserklärung ist und aufgezeigt, welche Regelungen zum Speichern und Archivieren beachtet werden sollten. Des Weiteren wird gezeigt, wie man durch Pseudonymisierung und Anonymisierung Daten minimieren kann. Es handelt sich hierbei ausdrücklich nicht um ein rechtssicheres Dokument, sondern lediglich um eine Empfehlung. Bei detaillierten Fragen bitten wir Sie, Kontakt zur Rechtsabteilung oder der/dem Datenschutzbeauftragten Ihrer Hochschule aufzunehmen.

Was sind personenbezogene Daten?

Die rechtliche Basis für den Umgang mit personenbezogenen Daten bildet die Datenschutzgrundverordnung (DSGVO) der Europäischen Union, welche durch das Bundesdatenschutzgesetz (BDSG) sowie die Landesdatenschutzgesetze, wie bspw. das Thüringer Datenschutzgesetz (ThürDSG), ergänzt bzw. konkretisiert wird. Personenbezogene Daten sind nach Art. 4 Abs. 1 der DSGVO „alle Informationen, die sich auf eine identifizierte oder identifizierbare natürliche Person beziehen. Als identifizierbar wird eine natürliche Person angesehen, wenn diese direkt oder indirekt mittels Zuordnung durch Namen, Standortdaten oder ähnliches erkannt werden kann.“ Dazu zählen Merkmale, „die Ausdruck der physischen, physiologischen, genetischen, psychischen, wirtschaftlichen, kulturellen oder sozialen Identität der natürlichen Person sind“. Art. 9 der DSGVO definiert zusätzlich besondere Kategorien personenbezogener Daten, die häufig umgangssprachlich als sensible personenbezogene Daten bezeichnet werden. An die Verarbeitung solcher Daten werden höhere technische und rechtliche Anforderungen gestellt.

Beispiele für betroffene Daten nach DSGVO:

Personenbezogene Daten

- Name
- Alter
- Schulbesuch
- Adresse
- Online-Kennung
- Telefonnummer
- Arbeitgeber*in und Arbeitsstelle
- Aktivitäten, Hobbies und Verhaltensmuster
- Fotos

Sensible personenbezogene Daten

- rassische und ethnische Herkunft
- politische Meinungen
- religiöse oder weltanschauliche Überzeugungen
- Gewerkschaftszugehörigkeit, Mitgliedschaft von Vereinen bzw. Organisationen
- genetische/biometrische Daten
- Gesundheitsdaten
- Sexualleben bzw. sexuelle Orientierung
- strafrechtliche Verurteilungen und Straftaten[1]

[1] nach Art. 10 DSGVO Verarbeitung nur unter behördlicher Aufsicht

Personenbezogene Daten -
Definition, Speicherung und Minimierung

FDM THÜRINGEN

Grundlage der Speicherung und die Einwilligungserklärung

Grundlegend dürfen nach DSGVO Art. 6 personenbezogene Daten erhoben und verarbeitet werden, wenn u.a. die betroffene Person in die Datenverarbeitung eingewilligt hat, die Datenverarbeitung zur Erfüllung eines Vertrages oder aufgrund rechtlicher Verpflichtungen erforderlich ist, oder berechtigte Forschungsinteressen des Verantwortlichen vorliegen. Allerdings handelt es sich bei letzterem um eine stets mit Rechtsunsicherheit verbundene Wertungsfrage, bei der die Forschungsinteressen die Grundrechte und Grundfreiheiten der betroffenen Person überwiegen müssen.

Um rechtliche Probleme zu vermeiden und die Transparenz des Forschungsvorhabens zu gewährleisten, ist es empfehlenswert eine Einwilligungserklärung der betroffenen Person für die Erhebung, Speicherung und ggfs. Veröffentlichung der Daten einzuholen. Unter dieser Einwilligung versteht man eine freiwillige, informierte und unmissverständliche Willensbekundung (Art. 4 Abs. 11 DSGVO). Auch wenn die Einwilligung mündlich erfolgen kann, ist es empfehlenswert diese schriftlich einzuholen, um zukünftige Konflikte zu vermeiden. Die Einwilligungserklärung sollte nach Art. 13 DSGVO folgende Informationen enthalten:

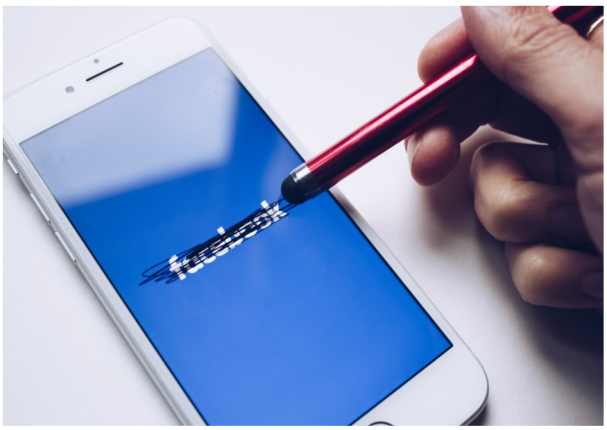
- Namen der Kontaktpersonen des Projekts (ggf. der/des Datenschutzbeauftragten)
- Zwecke der Verarbeitung der Daten bzw. die Rechtsgrundlage dafür
- Informationen zu Empfänger*innen der Daten (Personen, Organisationen, Länder, ...)
- Informationen dazu (bzw. Kriterien dafür), wie lange die Daten gespeichert werden
- Auskunft, welche Daten bei nicht betroffenen Personen verarbeitet werden und aus welchen Quellen sie stammen (Art. 14)

- Recht auf Auskunft (Art. 15), Berichtigung (Art. 16), Löschung (Art. 17), Einschränkung der Verarbeitung (Art. 18), Datenübertragung (Art. 20), Widerspruch (Art. 21), Verbot von automatisierter Entscheidungen und Profiling (Art. 22), Beschwerde und Widerruf der Einwilligung

Nicht alle Punkte müssen festgehalten werden, wenn diese entweder nicht für das Forschungsvorhaben zutreffen oder die Erhebung der Daten negativ beeinflussen würden (durch z.B. Bias der Teilnehmenden bei Beantwortung von Fragen). Selbstformulierte oder vorgegebene Mustererklärungen können verwendet, sollten aber von dem/der lokalen Datenschutzbeauftragten geprüft werden.

Die Informationspflicht an betroffene Personen über die Sammlung ihrer personenbezogenen Daten kann entfallen, wenn sich dies bspw. als unmöglich erweist oder einen unverhältnismäßigen Aufwand erfordert (Art. 14 Abs. 5b DSGVO und §27 BDSG). In dem Falle müssen allerdings technisch-organisatorische Maßnahmen zur Datenminimierung ergriffen werden (Art. 89 Abs 1. DSGVO).

Personenbezogene Daten -
Definition, Speicherung und Minimierung



Da personenbezogene Daten einer Zweckbindung und Speicherbegrenzung unterliegen (Art. 5 Abs. 1 DSGVO) dürfen diese nur solange gespeichert werden, wie dies für die Erfüllung des (Forschungs-)Zwecks erforderlich ist. Ist der Zweck durch das Ende des Forschungsprojekts erfüllt oder widerruft eine betroffene Person ihre ursprüngliche Einwilligung zur Datenverarbeitung, müssen nach Art. 17 DSGVO die personenbezogenen Daten gelöscht werden. Zur Einhaltung guter wissenschaftlicher Praxis ist eine Archivierung personenbezogener Daten für bis zu 10 Jahre jedoch weiterhin zulässig, solang geeignete Sicherungsmaßnahmen getroffen werden[1].

Personenbezogene Daten müssen sowohl während der Projektlaufzeit als auch gegebenenfalls nach dem Projektende, wenn diese weiterhin gespeichert werden müssen, durch organisatorische und technische Maßnahmen angemessen gesichert werden. Hierzu gehört es in der Regel, die Daten in einem sicheren und abschließbaren Raum aufzubewahren. Es empfiehlt sich außerdem, die Daten zu verschlüsseln und mit einem sicheren Passwort zu versehen. Das Passwort sollte dabei idealerweise zufällig generiert, dokumentiert und alle paar Jahre geändert werden. Damit der Zugang zu den Daten nicht durch eine Verkettung unglücklicher Umstände verloren geht, sollten immer zwei oder mehr Personen Zugang zu den Daten haben. Empfohlen ist es hierbei, eine/einen Sicherheitsbeauftragten zu bestimmen, der die Zugänge zu den Daten sowohl während der Projektzeit als auch nach dem Ende des Projekts organisiert und verwaltet.

Datenminimierung: Pseudonymisierung und Anonymisierung

Personenbezogene Daten sollten nach der Erhebung bereits zu einem möglichst frühen Zeitpunkt des Forschungsprojekts minimiert werden, soweit dies möglich ist und die Analyse nicht einschränkt. Dadurch können sie offener gehandhabt werden und unterliegen ebenfalls bei der Überführung in ein Archiv oder bei einer Veröffentlichung weniger Einschränkungen. Im Allgemeinen unterscheidet man bei der Datenminimierung zwischen der Pseudonymisierung und der Anonymisierung.

Pseudonymisierung

Nach Art. 4 Abs. 5 des DSGVO ist „Pseudonymisierung die Verarbeitung personenbezogener Daten in einer Weise, dass die personenbezogenen Daten ohne Hinzuziehung zusätzlicher Informationen nicht mehr einer spezifischen betroffenen Person zugeordnet werden können. Diese zusätzlichen Informationen müssen gesondert aufbewahrt werden. Außerdem sind technische und organisatorische Maßnahmen notwendig, die gewährleisten, dass die personenbezogenen Daten nicht einer identifizierten oder identifizierbaren natürlichen Person zugewiesen werden“.

Das heißt, unter Pseudonymisierung versteht man das Ersetzen von Namen oder anderen Identifikationsmerkmalen durch einen Code, der erst durch Zuhilfenahme von gesicherten Informationen wieder Rückschlüsse auf die eigentliche Person zulässt. Inwiefern das Pseudonym selber geschützt werden muss, hängt vom Aufwand zur Identifizierung der Person ab. Dies können objektive Faktoren sein, wie benötigte Kosten, Zeit, Technologie und Entwicklung.

> **Beispiel:** Wird eine zu bewertende Prüfung durchgeführt, können Teilnehmende einen erdachten Namen (Pseudonym) angeben. Dieser wird bei der Veröffentlichung in Zusammenhang mit dem einzelnen Ergebnis verwendet, sodass nur die/der Prüfende und die teilnehmende Person den Zusammenhang kennen.

[1] Wirth, Thomas (2020), Die Pflicht zur Löschung von Forschungsdaten - Urheber- und Datenschutzrecht im Widerspruch zu den Erfordernissen guter wissenschaftlicher Praxis?, Zeitschrift für Urheber- und Medienrecht (ZUM) 64(8/9), 585-592

Personenbezogene Daten -
Definition, Speicherung und Minimierung

Anonymisierung

Anonymisierte Daten beinhalten keinen Personenbezug, weshalb diese nach Erwägungsgrund 26 der DSGVO nicht mehr der DSGVO unterliegen und auch keine Definition dieser Daten dort formuliert ist. Allgemein gelten personenbezogene Daten als anonymisiert, wenn Personen über die Daten nicht mehr identifiziert und auch keine Rückschlüsse mehr über zu schützende Daten (wie bei der Pseudonymisierung) gezogen werden können. Der Vorgang der Anonymisierung ist also irreversibel.

In der Praxis werden personenbezogene Daten anonymisiert indem der eindeutige Identifikator gelöscht wird (wie beispielsweise der Name der Person oder ein anderer Schlüsselwert) und bestimmte Werte nicht mehr als eine Datenreihe gespeichert werden, wenn dies für die Analyse nicht mehr nötig ist (wie Ort, Zeit und Tätigkeit, welche zusammen wieder eine Person erkennbar machen können). Ein anderes Vorgehen ist die Aggregation von Daten. Hierbei werden Datensätze so gruppiert, dass sie ebenfalls keinen Rückschluss mehr auf die einzelne Person zulassen. Dies kann durch Runden oder Generalisierung von Werten erfolgen, indem man z.B. verschiedene Altersangaben in Altersgruppen zusammenfasst. Eine weitere Methode wäre die Maskierung, bei der die Originalwerte mit zufälligen (aber technisch validen) neuen Werten besetzt werden.

> **Beispiel:** Geheime Wahlen beruhen auf dem Prinzip der Anonymisierung, d.h. die Ergebnisse werden gesammelt und ausgewertet, aber es ist nicht möglich auf das Wahlverhalten einzelner Personen Rückschlüsse zu ziehen. In Datenbanken erfolgt die Anonymisierung im Allgemeinen dadurch, dass der verbindende Schlüssel zwischen den Tabellen (z.B. eine Kundennummer) gelöscht wird. Die restlichen Daten sind dadurch zusammenhangslos und können nicht mehr einer spezifischen Person zugeordnet werden.

Welche Maßnahmen in Bezug auf persönliche Daten angemessen oder gar erforderlich sind, sollte im Einzelfall mit der/dem lokalen Datenschutzbeauftragten geklärt werden.

Wurde jeglicher Bezug zur einzelnen Person aus den Forschungsdaten entfernt und diese damit anonymisiert, so können die Daten und die daraus entstanden Ergebnisse im Regelfall ohne weitere Einschränkungen veröffentlicht werden, wenn keine Interessen Dritter dagegensprechen (wie durch Lizenzen oder Auftraggeber).

Diese Handreichung bietet einen Einblick zum Umgang mit personenbezogenen Forschungsdaten. Sollten Sie weitere Fragen zu dem Thema haben, können Sie Ihre Rechtsabteilung, die lokal beauftragte Person zum Datenschutz oder unser Netzwerk unter https://forschungsdaten-thueringen.de/kontakt.html kontaktieren.