Der Landesbeauftragte für
den DATENSCHUTZ und die
INFORMATIONSFREIHEIT
Rheinland-Pfalz

Forschungsdatenschutz – Was muss ich als ForscherIn bei meinem Projekt beachten?

JGU FB 02 Sozialwissenschaften
Georg-Forster-Gebäude, Raum 01-711

MICHAEL SMOLLE
02. FEBRUAR 2026

Forschungsdatenschutz – Was muss ich als ForscherIn bei meinem Projekt beachten?

I. Theorie und rechtliche Grundlagen

II. Aus der aufsichtsbehördlichen Praxis

I. Theorie und rechtliche Grundlagen

1. Worum geht es überhaupt beim Datenschutz?

Im Zusammenhang mit einer von der Bundesregierung in den 1980ziger Jahren geplanten Volkszählung (Zensus) hat sich das Bundesverfassungsgericht (BVerfG) mit den Voraussetzungen für die Datenerhebung zu statistischen Zwecken beschäftigt und ein Grundrecht auf informationelle Selbstbestimmung jeder natürlichen Person im sog. „Volkszählungsurteil“ vom 15.12.1983 entwickelt.

Das Grundrecht ist im Text des GG nicht ausdrücklich beschrieben, sondern wurde vom BVerfG aus der allgemeinen Handlungsfreiheit (Art. 2 Abs. 1 GG) und der Menschenwürde (Art. 1 Abs. 1 GG) abgeleitet.

Als Anlass dafür sah das BVerfG die zunehmende automatisierte Datenverarbeitung mit Computern und überhaupt moderner Technologie im Verhältnis zur bis dahin überwiegenden Speicherung in Papierakten und Karteien.

Dadurch fällt es schwerer zu wissen, wer welche Daten von einem selbst wann und wofür verarbeitet.

BVerfGE 65, 1 (43)

In einer Zeit der Digitalisierung ist die Gewährleistung des Grundrechts von immer größerer Bedeutung!

2. Was wird von diesem Grundrecht geschützt?

Das Recht jedes einzelnen, selbst über die Verarbeitung und Preisgabe seiner personenbezogenen Daten zu entscheiden. Das Datenschutzrecht schützt also nicht die personenbezogenen Daten an sich, sondern die natürlichen Personen, über welche die Daten Informationen enthalten.

Datenschutz kann auch als Schutz der Menschen vor Schäden oder Beeinträchtigungen definiert werden, die aus der Ansammlung und missbräuchlichen Verwendung von personenbezogenen Daten entstehen können.

In der Fachsprache wird das Grundrecht als „Verbot mit Erlaubnisvorbehalt“ (Art. 6 Abs. 1 Datenschutz-Grundverordnung, DS-GVO) bezeichnet. Das bedeutet, dass die Verarbeitung von personenbezogenen Daten grundsätzlich verboten ist, wenn sie nicht durch die betroffene Person oder sonst durch Gesetz erlaubt ist.

Anders ausgedrückt sind personenbezogene Daten kein frei zugängliches Informationsmaterial, sondern jeder Zugriff ist eine Ausnahme, die begründet werden muss.

Denn jede Verarbeitung personenbezogener Daten stellt zunächst einmal ein Risiko für die Rechte und Freiheiten einer natürlichen Person dar.

BVerfGE 65, 1 (42, 44)

3. Erlaubnisse für die Verarbeitung personenbezogener Daten

Art. 6 Abs. 1 S. 1 DS-GVO zählt solche Erlaubnisse auf.

Für Wissenschaft und Forschung besonders relevant sind

- Einwilligung – lit. a
- Wahrnehmung einer Aufgabe im öffentlichen Interesse – lit. e

also Datenverarbeitung auf der Grundlage einer Einwilligung oder eines Gesetzes.

Beispiele für Wissenschaftsklauseln bzw. Rechtsvorschriften zur Verarbeitung personenbezogener Gesundheitsdaten zum Zweck der Wissenschaft und Forschung sind Art. 6 Abs. 1 S. 1 lit. e i.V.m. Abs. 2, 3 i.V.m.

- § 22 Landesdatenschutzgesetz (LDSG) RP
- § 37 Landeskrankenhausgesetz (LKG) RP
- § 12 Landeskrebsregistergesetze (LKRG) RP

Bundesgesetze zur Forschung mit Gesundheitsdaten
- § 8 Bundeskrebsregistergesetz (BKRG)
- § 6 Gesundheitsdatennutzungsgesetz (GDNG)

Art. 40 Abs. 4 Digital Services Act (DAS) Zugang zu nicht-öffentlichen Daten der großen Online-Plattformen / -Suchmaschinen

4. Was versteht man unter „personenbezogenen Daten“ und deren „Verarbeitung“?

Die maßgeblichen Definitionen sind zu finden in der für die Mitgliedstaaten der EU unmittelbar geltenden DS-GVO (Verordnung des Europäischen Parlaments und des Rates vom 27.04.2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten (……).

Laut Artikel 4 Nr. 1 und Nr. 2 dieser Verordnung bezeichnet der Ausdruck

„personenbezogene Daten“ alle Informationen, die sich auf eine identifizierte oder identifizierbare natürliche Person (betroffene Person) beziehen; als identifizierbar wird eine natürliche Person angesehen, die direkt oder indirekt, insbesondere mittels Zuordnung zu einer Kennung wie einem Namen, zu einer Kennnummer, zu Standortdaten, zu einer Online-Kennung oder zu einem oder mehreren besonderen Merkmalen, die Ausdruck der physischen, physiologischen, genetischen, psychischen, wirtschaftlichen, kulturellen oder sozialen Identität dieser natürlichen Person sind, identifiziert werden kann;

„Verarbeitung“ jeden …. Vorgang …. im Zusammenhang mit personenbezogenen Daten wie das Erheben, die Speicherung, die Übermittlung, die Verknüpfung, das Löschen u.a.“

5. In welche verschiedenen Kategorien können personenbezogene Daten unterschieden werden?
5.1 Direkt und indirekt identifizierende Daten

Personen können entweder direkt oder indirekt identifiziert werden. Für eine direkte Identifizierung sind eindeutige (bestimmte bzw. personenbezogene) Merkmale (wie Name, Adresse, Geburtsdatum, Sozialversicherungsnummer) nötig.

Personen können aber auch indirekt auf der Basis mehrerer (bestimmbarer bzw. personenbeziehbarer) Merkmale, die für sich genommen keine eindeutige Zuordnung zulassen, identifiziert werden.

Z. B. kann eine genaue und seltenere Berufsbezeichnung in Verbindung mit der Angabe des Geschlechts und des Wohnortes zur Identifizierung einer Person führen.

Identifizierende Daten (IDAT) wie Name, Adresse, Geburtsdatum werden zur Kommunikation, zur Erstellung von Zuordnungslisten oder eindeutigen Pseudonymen bei der Durchführung von Längsschnittstudien verwendet.

IDAT sind von den eigentlichen Forschungsdaten, die z.B. mit einem Fragebogen erhoben werden, getrennt zu speichern.

5. In welche verschiedenen Kategorien können personenbezogene Daten unterschieden werden?
5.1 Direkt und indirekt identifizierende Daten

Erstes Beispiel:

Eine Schule möchte die Auswirkungen des Projekts „Bewegte Pause“ auf die Schüler:innen überprüfen. Alle sollen einen Fragebogen ausfüllen, aber auf keinen Fall den Namen angeben und den ausgefüllten Bogen in eine „Wahlurne“ einwerfen. Allerdings wird im Fragebogen u.a. nach der Klassenbezeichnung, dem Geschlecht, dem Gewicht und der Körpergröße gefragt.

Für die Evaluation von Unterrichtsmethoden werden mit einem Fragebogen zwar keine Namen, aber Daten wie Klassenstufe, Herkunftsland, Muttersprache oder Aufenthaltsdauer in Deutschland abgefragt.

Was ist wenn, eine Lehrkraft der teilnehmenden Schule das Vorhaben durchführt und die ausgefüllten Fragebögen auswertet?

5. In welche verschiedenen Kategorien können personenbezogene Daten unterschieden werden?
5.1 Direkt und indirekt identifizierende Daten

Zweites Beispiel:

In einer Gemeinde mit weniger als 500 Einwohnern wird zur Vorbereitung auf die Dorferneuerung mit einem Fragebogen der Name der Wohnstraße, die in einem Haushalt lebende Anzahl von Personen sowie deren Alter und ein in der Gemeinde ausgeübtes Gewerbe (Schreiner, Winzer, Landwirt) abgefragt.

Der bzw. die Ortsbürgermeister:in wertet die ausgefüllten Fragebögen aus.

Damit war es im Hinblick auf die Größe und Einwohnerzahl der Gemeinde alles andere als ausgeschlossen, dass auf der Basis dieser Merkmale, die für sich genommen keine eindeutige Zuordnung zulassen, und wahrscheinlich vorhandenem Zusatzwissen auf eine bestimmte Person geschlossen und ihr die Angaben im Fragebogen zugerechnet hätten werden können.

5. In welche verschiedenen Kategorien können personenbezogene Daten unterschieden werden?
5.2 Pseudonyme Daten

Die DS-GVO beschreibt „Pseudonymisierung“ in Art. 4 Nr. 5 DS-GVO folgendermaßen:

die Verarbeitung personenbezogener Daten in einer Weise, dass die personenbezogenen Daten ohne Hinzuziehung zusätzlicher Informationen nicht mehr einer spezifischen betroffenen Person zugeordnet werden können, sofern diese zusätzlichen Informationen **gesondert aufbewahrt** werden und technischen und organisatorischen Maßnahmen unterliegen, die gewährleisten, dass die personenbezogenen Daten nicht einer identifizierten oder identifizierbaren natürlichen Person zugewiesen werden.

Bei pseudonymen Daten handelt es sich also noch um personenbezogene Daten.

5. In welche verschiedenen Kategorien können personenbezogene Daten unterschieden werden?
5.2 Pseudonyme Daten

Pseudonymisierung stellt eine Konkretisierung des Grundsatzes der Datenvermeidung und Datensparsamkeit (Art. 5 Abs. 1 lit. c DS-GVO) als einem vorrangigen Ziel des Datenschutzes dar, wonach keine oder so wenig personenbezogene Daten wie möglich zu verarbeiten sind.

Auf das notwendige Maß ist die Datenverarbeitung dann beschränkt, wenn nur solche Daten verwendet werden, ohne die der Zweck der Verarbeitung nicht erreicht werden kann.

Wenn für einen bestimmten Zweck beispielsweise die Verarbeitung des Lebensalters ausreicht, darf nicht das vollständige Geburtsdatum verarbeitet werden.

Dieser Grundsatz erstreckt sich aber nicht nur auf die Menge der verarbeiteten Daten, sondern auch auf die Dauer der Zugänglichkeit.

5. In welche verschiedenen Kategorien können personenbezogene Daten unterschieden werden?
5.2 Pseudonyme Daten

Müssen zunächst personenbezogene Daten erhoben werden, kann dem oben genannten Grundsatz der Datensparsamkeit und dem Recht auf informationelle Selbstbestimmung aber zu einem späteren Zeitpunkt mit einer Anonymisierung bzw. einer Pseudonymisierung der erhobenen Daten Rechnung getragen werden.

Daraus ergibt sich die Pflicht, in jeder Phase eines Forschungsvorhabens zu prüfen, ob unter Berücksichtigung des Forschungszwecks eine Veränderung der Einzelangaben in der Weise möglich ist, dass diese einer bestimmten Person nicht mehr zugeordnet werden können. Dies kann der Fall sein nach dem Abschluss der Prüfung der Einzelangaben auf Plausibilität.

Art. 89 Abs. 1 DS-GVO

5. In welche verschiedenen Kategorien können personenbezogene Daten unterschieden werden?
5.2 Pseudonyme Daten

Pseudonymisierung

- hat das Ziel, die unmittelbare Kenntnis der vollen Identität des Betroffenen während solcher Verarbeitungsvorgänge, bei denen ein Personenbezug nicht erforderlich ist, auszuschließen.
- verringert die Verknüpfbarkeit eines Datenbestands mit der Identität einer Person und stellt somit eine sinnvolle Sicherheitsmaßnahme dar.
- kann das Ersetzen des Namens und anderer direkter Identifikationsmerkmale durch ein Kennzeichen zu dem Zweck bedeuten, die Bestimmung einer Person auszuschließen oder wesentlich zu erschweren.

Erlaubnis

5. In welche verschiedenen Kategorien können personenbezogene Daten unterschieden werden?
5.2 Pseudonyme Daten

Im Falle der Verwendung einer **Zuordnungs- oder Referenzliste** werden die eine Person unmittelbar identifizierenden Daten durch eine für das Einzelvorhaben zu bildende Zuordnungsvorschrift derart verändert, dass eine mit dem Pseudonym verknüpfte Forschungsdatensatz nur mit Kenntnis dieser Zuordnungsvorschrift wieder einer natürlichen Person zugeordnet werden kann.

So wird beispielsweise in einer Tabelle dem Namen eine Forschungs-ID zugeordnet.

Idealerweise werden die Aufgaben auf verschiedene Verantwortliche verteilt. Von einem wird die Pseudonymisierung durchgeführt, ein anderer verwahrt die Zuordnungsregel und wiederum andere dürfen die pseudonymen Forschungsdaten verarbeiten. Erfolgt die Pseudonymisierung zudem von einem vertrauenswürdigen, unabhängigen Dritten (Datentreuhänder oder Vertrauensstelle), wird das Risiko einer Re-Identifizierung weiter gemindert.

Die Berechtigung für den Zugriff auf eine Zuordnungsregel ist festzulegen und technisch und organisatorisch zu gewährleisten. Vorstellbar wäre die Speicherung auf einem Rechner, zu dem nur der Datenschutzbeauftragte eines Unternehmens oder einer Forschungseinrichtung passwortgeschützt Zugriff nehmen kann.

Die Beschäftigten bzw. die Wissenschaftler arbeiten selbst nur mit einem Datensatz, der keine identifizierenden Daten enthält.

5. In welche verschiedenen Kategorien können personenbezogene Daten unterschieden werden?
5.2 Pseudonyme Daten

Folgendes Anwendungsbeispiel:

Eine Schule nimmt seit mehreren Jahren an einer wissenschaftlichen Studie teil, für die Informationen bzw. Einzelangaben über den Verlauf der Schullaufbahnen von Schüler:innen von der 5. bis zur 10. Klassenstufe zwecks wissenschaftlicher Analyse und Grundlage für bildungspolitische Entscheidungen erhoben und verarbeitet werden sollen.

In jedem Schuljahr müssen die teilnehmenden Schüler:innen mindestens einen Fragebogen ausfüllen. Der Fragebogen enthält nie den Namen, aber für eine bestimmte Schülerin bzw. einen bestimmten Schüler immer dieselbe Nummer, damit die Informationen über die Jahre immer derselben Person zugeordnet werden können.

Dafür wird in der Schule eine sog. Referenzliste geführt, in der die Schüler:innen mit Namen und Fragebogennummer aufgezählt werden. Ein ausgefüllter Fragebogen kann über die Nummer als Pseudonym nur mit Kenntnis dieser Liste wieder einer natürlichen Person zugeordnet werden. Für die Wissenschaftler enthalten die ausgefüllten Fragebögen somit nur pseudonyme Daten.

Was ist die Folge der Löschung einer Referenzliste?

6. Was sind anonyme oder anonymisierte Daten?

Einzelangaben können auch schon anonym, also ohne Bezug zu einer bestimmten oder bestimmbaren Person, erhoben werden

Der Ausdruck „Anonymisieren“ bezeichnet das Verändern personenbezogen erhobener Daten dergestalt, dass die Einzelangaben über persönliche oder sachliche Verhältnisse nicht mehr oder nur mit einem unverhältnismäßigen Aufwand an Zeit, Kosten und Arbeitskraft einer bestimmten oder bestimmbaren natürlichen Person zugeordnet werden können.

Um das zu erreichen wird der Bezug zu einer Person z.B. dadurch aufgehoben, dass in einem Datensatz die identifizierenden Daten gelöscht werden.

Anonym erhobene bzw. im Verlauf der weiteren Datenverarbeitung anonymisierte Daten unterliegen nicht dem Datenschutzrecht.

Datenschutzrechtliche Nutzungsbeschränkungen bestehen für solche Daten nicht (vgl. Erwägungsgründe 26, 28 zur DS-GVO).

6. Was sind anonyme oder anonymisierte Daten?

Das Weglassen des Namens oder der Anschrift als unmittelbar identifizierende Daten und die lediglich unvollständige Speicherung der IP-Adresse (sog. personenbeziehbares Datum) reicht aber mitunter nicht aus, um eine Anonymität zu erreichen.

Die Zuordnung einer Einzelangabe zu einer betroffenen Person muss aber auch nicht schlechthin ausgeschlossen sein, sondern es genügt für eine erfolgreiche Anonymisierung, wenn eine Zuordnung bzw. eine Identifizierung einer natürlichen Person anhand der vorhandenen Daten nach der Lebenserfahrung nicht zu erwarten ist.

Zur Beantwortung der Frage, ob trotz erhobener personenbeziehbarer Daten – soziodemografische Daten wie Alter, Geschlecht, Herkunft - von Anonymität auszugehen ist, muss eine Analyse des Risikos für eine Re-Identifizierung im Einzelfall erfolgen.

(Faktische) Anonymität ist letztlich eine Frage der Wahrscheinlichkeit.

6. Was sind anonyme oder anonymisierte Daten?

Es gibt **3** unterschiedliche Begriffe von Anonymität:

Formal – der Prozess der Anonymisierung beschränkt sich auf die Löschung von unmittelbar identifizierenden Daten, wie z.B. dem Namen, der Anschrift oder dem Geburtsdatum.

Faktisch – dem liegt der Ansatz zugrunde, dass der der Begriff des Personenbezuges relativ ist, d.h. es ist auf das konkrete Wissen des Verantwortlichen abzustellen. Die Zuordenbarkeit ist von den individuellen Fähigkeiten dieser Stelle abhängig. Nur wenn die Zuordnung der Daten zu einer bestimmten Person mit den dort zur Verfügung stehenden Hilfsmitteln erfolgen kann, soll ein Personenbezug bestehen.

Die Zuordnung einer Einzelangabe zu einer betroffenen Person muss nicht schlechthin ausgeschlossen sein, sondern es genügt für eine erfolgreiche Anonymisierung, wenn eine Zuordnung nach der Lebenserfahrung nicht zu erwarten ist.

Absolut – Hier wird nicht auf die tatsächlichen, individuellen Möglichkeiten des Verantwortlichen abgestellt, sondern vielmehr objektiv auf die generell verfügbaren Verknüpfungstechniken oder das in Theorie verfügbare Zusatzwissen, um den Bezug herstellen zu können. Es reicht jede theoretische, von den tatsächlichen Möglichkeiten des Verantwortlichen losgelöste Verknüpfung zwischen Person und Datum aus.

Der Ansatz faktischer Anonymität setzt sich immer mehr durch – wissenschaftsfreundlich !!

6. Was sind anonyme oder anonymisierte Daten?

Faktische Anonymität ist letztlich eine Frage der Wahrscheinlichkeit

Zur Beantwortung der Frage, ob trotz erhobener personenbeziehbarer Daten von anonymen Daten auszugehen ist, muss eine Analyse des Risikos für eine Re-Identifizierung im Einzelfall erfolgen.

Das Re-Identifizierungsrisiko bezeichnet die Eintrittswahrscheinlichkeit einer möglichen De-Anonymisierung von faktisch anonymisierten Daten unter Berücksichtigung der aus einer De-Anonymisierung möglicherweise entstehenden Folgen für die betroffene Person.

Werden zu einem wissenschaftlichen Zweck Angaben zu Geschlecht, Alter und Fächerkombination von den Lehrkräften einer überschaubaren Zahl von Grundschulen erhoben, besteht für die Re-Identifizierung einer männlichen Lehrkraft mittleren Alters wegen über die Schul-Homepage einfach zu erlangendem Zusatzwissen wohl ein hohes Risiko.

Faktische Anonymität ist ein dynamischer Zustand und Bedarf eines fortlaufenden Prozesses der Prüfung der Wirksamkeit der Anonymität!

6. Was sind anonyme oder anonymisierte Daten?

Kurzer Exkurs

Davon abzugrenzen sind aggregierte Daten.

Solche zusammengefassten Daten, wie z.B. die sich aus einer statistischen Erhebung ergebenden Summenangaben zu den einzelnen Erhebungsmerkmalen, enthalten keine Einzelangaben zu einer Person mehr.

Anonyme Daten enthalten dagegen mindestens eine Einzelangabe über eine Person, ohne dass die Person allerdings bekannt ist.

6. Was sind anonyme oder anonymisierte Daten?

Grundlegende Verfahren zur Anonymisierung

Generalisierung - ersetzen eines genauen Datums durch einen weniger spezifischen Wert, z.B.

- Vollständiges Geburtsdatum durch Geburtsjahr
- Geburtsjahr durch Zeitraum Lebensalter 50 – 60 Jahre
- die Angabe einer Region statt einer Stadt oder eines Monats statt einer Woche
- …

Randomisierung – eine Reihe von Techniken, welche die Daten in einer Weise verfälschen, dass durch hinzugefügte Daten die direkte Verbindung zwischen Daten und Betroffenem entfernt wird.

Weiterführend beispielsweise Artikel-29-Datenschutzgruppe, WP 216

6. Was sind anonyme oder anonymisierte Daten?

Risiken für eine robuste Anonymisierung:

- Herausgreifen *oder singling out*, d. h. die Möglichkeit, in einem Datenbestand einige oder alle Datensätze zu isolieren, welche die Identifizierung einer Person ermöglichen

- Verknüpfbarkeit, d. h. die Fähigkeit, mindestens zwei Datensätze, welche dieselbe Person oder Personengruppe betreffen, zu verknüpfen (in derselben Datenbank oder in zwei verschiedenen Datenbanken).

- Inferenz, d. h. die Möglichkeit, den Wert eines Merkmals mit einer signifikanten Wahrscheinlichkeit von den Werten einer Reihe anderer Merkmale abzuleiten.

7. Fazit oder die vor Beginn eines Forschungsvorhabens zu stellende Frage

Schon das BVerfG hat in dem oben erwähnten „Volkszählungsurteil“ ausgeführt, dass immer zu prüfen ist, ob das Ziel einer Datenerhebung nicht auch durch anonyme bzw. anonymisierte Daten erreicht werden kann und wenn dies nicht möglich ist, die Verarbeitung personenbezogener Daten sich auf das erforderliche Minimum oder auf pseudonymisierte Daten zu beschränken hat.

Kann der Zweck der Forschung mit verhältnismäßigem Aufwand auch mit ausschließlich anonymen Daten erreicht werden?

Unterscheidung zwischen Querschnitt- und Längsschnittstudie mit mehreren Messzeitpunkten, Follow Up-Erhebung.

II. Aus der aufsichtsbehördlichen Praxis

1. Informierte Einwilligung

Grundlage für eine rechtmäßige Datenverarbeitung kann u.a. die Einwilligung einer betroffenen Person sein (Art. 6 Abs. 1 lit. a DS-GVO).

Die datenschutzrechtlichen Anforderungen an Einwilligungen ergeben sich insbesondere aus Art. 7 DS-GVO.

Kerninhalte einer Information finden sich auch in den Erwägungsgründen 32 und 42 zur DS-GVO.

Grundsätzlich gilt, dass für eine wirksame Einwilligung die betroffenen Personen über Rechte nach Art. 13 Abs. 2 lit. b bis d DS-GVO und über die wesentlichen Modalitäten der vorgesehenen Datenverarbeitung aufzuklären sind.

Nur wer hinreichend abschätzen kann, welche Folgen die Erteilung oder Nichterteilung einer Einwilligung für ihn haben wird, kann selbstbestimmt darüber entscheiden, ob er der Erhebung und Verwendung seiner Daten zustimmt.

1. Informierte Einwilligung

Zu u.a. folgenden Fragen und Aspekten müssen die Betroffenen im Einzelnen aufgeklärt werden, damit sie in der Lage sind abzuschätzen, in was sie mit der Teilnahme an einem Projekt im Hinblick auf die Datenverarbeitung einwilligen:

Hinweis, dass die Teilnahme und die Einwilligung in eine Datenverarbeitung freiwillig ist und jederzeit ohne nachteilige Folgen abgebrochen oder widerrufen werden kann.

Warum werden personenbezogene Daten für den Forschungszweck benötigt?

Welchem Zweck dient die Datenverarbeitung? Welche Instrumente werden eingesetzt (Fragebogen, Interview, Test, Video)?

Wer ist verantwortlich, wie sieht das Prozedere für die Durchführung eines Vorhabens aus?

Gibt es mehrere Messzeitpunkte?

Werden Datensätze aus Messzeitpunkten miteinander verknüpft? Wie?

1. Informierte Einwilligung

Wer ist verantwortlich, wie sieht das Prozedere für die Durchführung eines Vorhabens aus?

Wie wird im Laufe des Vorhabens ggf. die Anonymität gewährleistet?

Wer erhält ggf. Kenntnis bzw. Zugang zu personenbezogenen Daten?

Wie werden verarbeitete Daten vor dem Zugriff durch unbefugte Dritte geschützt?

Werden Daten an Dritte übermittelt oder veröffentlicht?

Wann werden die Daten gelöscht?

1. Informierte Einwilligung

Bei einer Online-Befragung könnte die Einwilligung der daran teilnehmenden Personen aber bei einer entsprechenden vorherigen Information in einem elektronischen Opt-in oder in dem Versenden eines ausgefüllten Fragebogens gesehen werden.

Aus Art.77 Abs.1 DSGVO ergibt sich, dass sich die betroffene Person an jede beliebige europäische Aufsichtsbehörde wenden kann.

Würde man nun nur eine oder zwei Aufsichtsbehörden gegenüber der betroffenen Person benennen, könnte man Gefahr laufen, dass diese fälschlicher Weise glaubt, sich nur an diese Behörde wenden zu können.

1. Informierte Einwilligung

EDSA Leitlinien 05/2020 zur Einwilligung gemäß Verordnung 2016/679
Version 1.1, angenommen am 4. Mai 2020

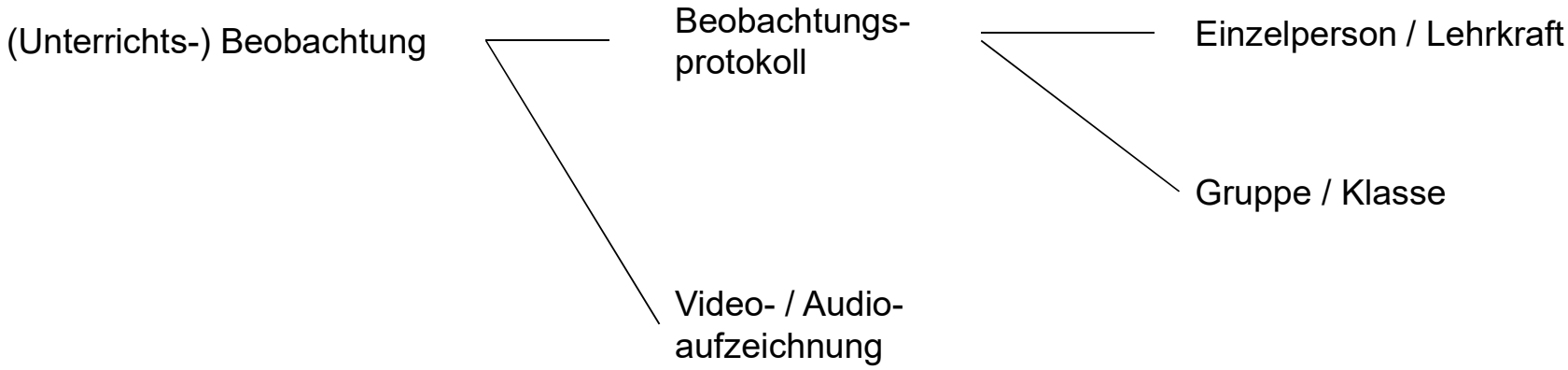
https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_202005_consent_de.pdf

Gem. § 67 Abs. 7 Schulgesetz Rheinland-Pfalz sind alle **wissenschaftlichen Untersuchungen an Schulen**, die nicht von den Schulen selbst, von den Schulbehörden oder den Schulträgern durchgeführt werden, durch die Aufsichts- und Dienstleistungsdirektion zu genehmigen.

https://add.rlp.de/themen/schule-und-bildung/schuelerinnen-und-eltern/wissenschaftliche-untersuchungen

2. Erhebung im Zuge qualitativer Forschung

Regelmäßig ergibt sich schon aufgrund des persönlichen Kontakts zu der befragten oder beobachteten Person (bei der Datenerhebung zunächst) eine Personenbeziehbarkeit der erhobenen Daten.



3. Datenübermittlung an Drittländer oder internationale Organisationen

Kapitel V der DS-GVO

Stichwort **„vergleichbares Datenschutzniveau“** Art. 44 S. 2 DS-GVO

Prüfung des „2-Stufen-Modells“ Art. 44 S. 1 DS-GVO:

Verarbeitungsgrundlage gemäß Art 6 Abs. 1 **und** Art. 44 ff. DS-GVO gefordert.

Angemessenheitsbeschluss der EU-Kommission gemäß Art. 45 DS-GVO für das Datenschutzniveau eines Drittlandes

Beispiel USA – EU-US Data Privacy Framework
lediglich **sektorale** Wirkung des Angemessenheitsbeschlusses !!

aktuelle Liste der Länder mit Angemessenheitsbeschluss finden Sie auf

https://commission.europa.eu/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en

3. Datenübermittlung an Drittländer oder internationale Organisationen

Kapitel V der DS-GVO

Alternative zu einem Angemessenheitsbeschluss - eine (zweite) Einwilligung gemäß Art. 49 Abs. 1 S. 1 lit. a DS-GVO

Dazu auch die

„Anwendungshinweise zu den Anforderungen an Datenübermittlungen an Drittländer im Rahmen der wissenschaftlichen Forschung zu medizinischen Zweck“

der Konferenz der unabhängigen Datenschutzaufsichtsbehörden des Bundes und der Länder DSK unter

https://www.datenschutzkonferenz-online.de/media/oh/20250917_DSK_OH_Datenuebermittlungen.pdf

Insgesamt zu dem Thema: Webinar #3: Biotechnologie und internationaler Datenschutz

https://www.datenschutz.rlp.de/themen/veranstaltungen/webinar-biotechnologie-goes-datenschutz

Weitere Alternative: Übermittlung von Daten entsprechend Entscheidung des EuGH vom 04.09.2025, Az. 413/23 P

„anonymisierende Wirkung von pseudonymen Daten“

4. Frage an Probanden nach strafbarem Verhalten

Verschiedene Herangehensweisen:

Anonym gestaltete Datenerhebung

Informierte Einwilligung

Benennung der Risiken, weil derzeit kein Forschungsgeheimnis mit Beschlagnahmeverbot für Strafverfolgungsbehörden und einem Zeugnisverweigerungsrecht für Forscher:innen existiert.

Im Falle einer Befragung von minderjährigen Schüler:innen – grds. Einwilligung seitens der Eltern bzw. der Erziehungsberechtigten

Gesetzliche Grundlage § 476 StPO Auskunft in Form der Überlassung von Kopien und Akteneinsicht für Forschungszwecke

5. Künstliche Intelligenz und Datenschutz

- Die Anwendbarkeit der DS-GVO bleibt von KI-VO unberührt, d.h.
- die DS-GVO findet auch Anwendung bei der Entwicklung und Verwendung von KI-Modellen !

  Daher:
  - ➢ DS-GVO enthält keine spezifischen Vorschriften für KI-Systeme
  - ➢ Prüfung zahlreicher rechtlicher Aspekte anhand der DS-GVO

    - Einschlägige Rechtsgrundlage für die Verarbeitung personenbezogener Daten
    - Festlegung der Verantwortlichkeit
    - Gewährleistung von Betroffenenrechten
    - DSFA
    - usw.

5. Künstliche Intelligenz und Datenschutz

Einwilligung?

- Jederzeit Widerruf möglich (Art. 7 Abs. 3 Satz 1 DSGVO) und in Folge dessen Recht auf Löschung: bei Trainingsdaten kaum möglich

- Daten Dritter aus dem Internet: von den i.d.R. unbekannten Personen kann kaum eine Einwilligung eingeholt werden.

- Trainingsdaten: Festlegung auf eindeutige und abgrenzbare Zwecke der Datenverarbeitung kann schwierig sein wegen autonomer Weiterentwicklung des Modells und damit verbundener kontinuierlicher Datenverarbeitung mit häufiger Änderung der Verarbeitungszwecke

- Informiertheit bei komplexen Systemen: umfassende Aufklärung über die Funktionsweise der eingesetzten KI möglich ? Komplexität der Logik auf ein verständliches Maß runterbrechen. Algorithmus als Geschäftsgeheimnis ?

Informationsmaterial und Quellen zu KI

Konferenz der unabhängigen Datenschutzaufsichtsbehörden des Bundes und der Länder (DSK):

- Orientierungshilfe Künstliche Intelligenz und Datenschutz, Vs. 1.0, 06.05.2024
- Orientierungshilfe zu empfohlenen technischen und organisatorischen Maßnahmen bei der Entwicklung und beim Betrieb von KI-Systemen, Vs. 1.0, 06/2025
- Orientierungshilfe zu datenschutzrechtlichen Besonderheiten generativer KI-Systeme mit RAG-Methode, Version 1.0, 10/2025

https://www.datenschutz.rlp.de/service/aktuelles/detail/orientierungshilfe-zu-ki-systemen-mit-retrieval-augmented-generation

15 Aspekte zum kontrollierten Umgang mit LLM-Chatbots

https://datenschutz-hamburg.de/news/checkliste-zum-einsatz-llm-basierter-chatbots

EDSA Stellungnahme zu gewissen Datenschutzaspekten der Verarbeitung personenbezogener Daten im Zusammenhang mit KI-Modellen (28/2024)

https://www.edpb.europa.eu/system/files/2025-05/edpb_opinion_202428_ai-models_de.pdf

Webinar #4: Datenschutz, KI und Biotechnologie
https://www.datenschutz.rlp.de/themen/veranstaltungen/webinar-biotechnologie-goes-datenschutz

6. Sonstiges

- Forscher programmieren selber eine App, die u.a. Gesundheits-Daten abfragt und die später kommerziell genutzt werden soll.

  „Kommt auf die konkrete Vorgehensweise an.“

- Forscher möchten für ihr Projekt einen Chatbot von einer externen Firma programmieren lassen, auf deren Website kein Impressum ist und unter „Datenschutz“ steht dort, dass sich an alle Datenschutz-Vorgaben gehalten wird. Die Chefs der Firma sind Israelis und die Programmierer alles Russen.

  „Finger weg?!“

FRAGEN

Materialien

Leitlinien 05/2020 zur Einwilligung gemäß Verordnung 2016/679 Version 1.1, angenommen am 4. Mai 2020

https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_202005_consent_de.pdf

Guidelines 01/2025 on Pseudonymisation des European Data Protection Board vom 16. Januar 2025

Leitlinien 07/2020 zu den Begriffen „Verantwortlicher“ und „Auftragsverarbeiter“ in der DSGVO
Version 2.0, angenommen am 7. Juli 2021

https://www.edpb.europa.eu/system/files/2023-10/edpb_guidelines_202007_controllerprocessor_final_de.pdf

Zur Zeit werden Leitlinien des Europäischen Datenschutzausschusses (EDSA) zur Anonymisierung personenbezogener Daten sowie zur Datenverarbeitung für Zwecke wissenschaftlicher Forschung erarbeitet.

https://www.datenschutz.rlp.de/themen/verarbeitung-forschungszwecke

mit Verlinkung auf das weitere Angebot der hessischen Datenschutzaufsichtsbehörde

VIELEN DANK FÜR IHRE AUFMERKSAMKEIT!

Der Landesbeauftragte für den
DATENSCHUTZ und die
INFORMATIONSFREIHEIT
Rheinland-Pfalz

Michael Smolle

Bereichsleiter Datenschutz 2
Kommunales; Hochschulen; Wissenschaft; Kultur

beim Landesbeauftragten für den Datenschutz
und die Informationsfreiheit Rheinland-Pfalz

Postanschrift:	Postfach 30 40 55020 Mainz
Büroanschrift:	Hintere Bleiche 34 55116 Mainz
Telefon:	+49 (6131) 8920-220
E-Mail:	m.smolle@datenschutz.rlp.de
Web:	www.datenschutz.rlp.de