IT-PLANUNGSRAT

Hamburg
28.04.2025

Handreichung Anonymisierung

Datenschutzrechtliche und technische Rahmenbedingungen bei der Anonymisierung im Kontext der Entwicklung und Nutzung von KI-Systemen

Version	Datum	Autor:in	Aktion
1.0.	28.04.2025	Frank Becker	Erstellung

Inhaltsverzeichnis

1 Einleitung

Diese Handreichung richtet sich an alle diejenigen, die in der öffentlichen Verwaltung tätig sind und dort mit dem Training von KI-Systemen, dem Datenschutz oder der Entwicklung von KI-Systemen (ggf. durch einen Dienstleister) betraut sind.
Die Handreichung soll einen knappen und strukturierten Überblick über die Grundlagen und Anforderungen an die Anonymisierung in technischer und datenschutzrechtlicher Hinsicht geben. Sie gibt Hilfestellung und einen Überblick über die Hintergründe und Möglichkeiten der Anonymisierung, kann aber nicht auf jeden Einzelfall eingehen. Ein Austausch mit Datenschutzreferent:innen oder eine Rückfrage an die Datenschutz-Aufsichtsbehörde bei komplexen Fragestellungen ist immer angeraten.

Anonymisierung hat im Datenschutzrecht eine **Schutzfunktion** und eine **Ermöglichungsfunktion**.

- **Schutzfunktion:** Die Verarbeitung personenbezogener Daten birgt Risiken für das Persönlichkeitsrecht der Betroffenen. Diese Risiken werden durch die Anonymisierung gesenkt. Denn anonymisierte Daten können nicht mehr mit den Betroffenen in Verbindung gebracht und somit nicht zu ihren Lasten missbraucht werden. Dies entspricht dem Schutzziel des Datenschutzrechts. Daher fordert das Datenschutzrecht, personenbezogene Daten zu löschen oder zu anonymisieren, sobald und soweit die Identifizierbarkeit der Betroffenen für die Zwecke der Verarbeitung nicht mehr erforderlich ist.[1]

[1] Vgl. Art. 5 Abs. 1 Buchst. e DSGVO.

- **Ermöglichungsfunktion:** Die Ermöglichungsfunktion der Anonymisierung liegt darin, dass sie in bestimmten Fällen eine beabsichtigte Verarbeitung überhaupt erst datenschutzrechtlich zulässig macht. So kann es sein, dass das Training einer KI mit personenbezogenen Daten unzulässig, mit anonymisierten Daten hingegen zulässig ist. Das Datenschutzrecht verlangt an verschiedenen Stellen Rücksicht auf die Betroffenen (z.B. beim Grundsatz der Datenminimierung und bei Zweckänderungen) und enthält die allgemeine Pflicht, den Risiken für die Betroffenen durch angemessene Maßnahmen zu begegnen.[2] Daher kann die datenschutzrechtliche Prüfung zu dem Ergebnis führen, dass nur mit dem durch die Anonymisierung bewirkten Schutz der Betroffenen die Verarbeitung (zum Beispiel das Training einer KI) rechtmäßig ist und nur so den mit der Verarbeitung verbundenen Risiken angemessen begegnet werden kann.

Anonymisierung kann rechtlich notwendig sein bei:

- Statistischen Auswertungen
- Erfüllung von Transparenzverpflichtungen
- Vorbereitung von Daten für die Weiterverwendung
- Der Entwicklung von KI-Systemen

Konkret tritt die **Anonymisierung als rechtliche Anforderung** unter anderem im Rahmen statistischer Auswertungen,[3] bei der Erfüllung von Transparenzpflichten[4] sowie bei der Vorbereitung von Daten für eine Weiterverwendung[5] auf und ist auch bei der Entwicklung von KI-Systemen[6] relevant.

Um Anonymisierungsprozesse schneller und effizienter zu gestalten, ihre Wirksamkeit zu gewährleisten und Standards und Best Practices zu etablieren, werden in mehreren Ländern der Bundesrepublik Deutschland derzeit **KI-Systeme zur Anonymisierung und Pseudonymisierung** personenbezogener Daten entwickelt. Die Einführung derartiger KI-Systeme wird

[2] Vgl. Art. 32 Abs. 1 DSGVO.
[3] Vgl. § 11 Absatz 2 HmbDSG.
[4] Vgl. § 4 Absatz 1 HmbTG.
[5] Vgl. Art. 5 Abs. 3 Buchst. a Ziff. i Daten-Governance-Rechtsakt (VO (EU) 2018/1714) und Erwägungsgrund 52 Open Data und PSI-Richtlinie (RL (EU) 2019/1024).
[6] Vgl. § 8 Absatz 1 und 2 ITEG SH und § 13 Abs. 2 HmbVwDiG.

allerdings von technischen und rechtlichen Hürden begleitet, die den Entwicklungsprozess verzögern.

2 Definitionen

2.1 Personenbezogene Daten

Personenbezogene Daten sind nach DSGVO alle Informationen, die sich auf eine identifizierte oder identifizierbare natürliche Person beziehen.[7] **Identifizierbar** ist eine natürliche Person dann, wenn sie direkt oder indirekt, insbesondere mittels Zuordnung zu einer Kennung wie einer Kennnummer oder einem besonderen Merkmal, das Ausdruck der Identität dieser natürlichen Person ist, identifiziert werden kann.

Maßstab der Identifizierbarkeit: Zur Beurteilung, ob eine natürliche Person identifizierbar ist, ist auf alle Mittel abzustellen, die von dem Verantwortlichen oder einer anderen Person nach allgemeinem Ermessen wahrscheinlich genutzt werden, um die natürliche Person direkt oder indirekt zu identifizieren.[8] Bei der Feststellung, ob Mittel nach allgemeinem Ermessen wahrscheinlich zur Identifizierung der natürlichen Person genutzt werden, sollten alle objektiven Faktoren, wie Kosten der Identifizierung und der dafür erforderliche Zeitaufwand, herangezogen werden, wobei die zum Zeitpunkt der Verarbeitung verfügbare Technologien zu berücksichtigen sind.[9]

Die folgenden Beispiele aus der **Rechtsprechung des EuGH** helfen, diesen abstrakten Maßstab zu konkretisieren:

- **Fahrzeugidentifikationsnummer (FIN):** Für den Personenbezug einer FIN kommt es nach Ansicht des EuGH darauf an, ob die Stelle, die die FIN verarbeitet, über Mittel verfügt, die es ihr ermöglichen, die FIN einer bestimmten Person zuzuordnen.[10] Der Personenbezug ist also **relativ** zu bestimmen, d.h. aus der Sicht der datenverarbeitenden Stelle.

[7] Vgl. Art. 4 Nr. 1 DSGVO.
[8] Vgl. Erwägungsgrund 26 Satz 3 DSGVO.
[9] Vgl. Erwägungsgrund 26 Satz 4 DSGVO.
[10] EuGH, Urt. v. 9.11.2023 – C-319/22, Rn. 46.

- **IP-Adressen:** Trotz der relativen Betrachtungsweise können Identifikationsmittel, die in den Händen Dritter liegen, bei der Bestimmung des Personenbezugs zu berücksichtigen sein. So ist nach der EuGH-Rechtsprechung zum Personenbezug von IP-Adressen für die Beurteilung der Identifizierbarkeit einer Person nicht allein auf die beim Verantwortlichen vorhandenen Informationen abzustellen.[11] Es genüge, wenn dem Verantwortlichen rechtliche Mittel zur Verfügung stehen, die es ihm erlauben, gegebenenfalls über eine andere Stelle die betroffene Person bestimmen zu lassen.[12] Von einer wahrscheinlichen Nutzung von Mitteln zur Identifizierung einer Person sei aber nicht auszugehen, wenn die Identifizierung **gesetzlich verboten** oder wegen **unverhältnismäßigen Aufwands an Zeit und Kosten** praktisch nicht durchführbar wäre.[13]
- **TC-String (bei Onlinewerbung):** Diesen relativen Maßstab zur Bestimmung des Personenbezugs einer Information abhängig von der Stelle, die die Information verarbeitet, wendete der EuGH auch in seiner Entscheidung zum sog. TC-String an. Der TC-String ist eine Zeichenfolge, die von einem Unternehmen namens IAB Europe generiert wird und Präferenzen eines Nutzers einer (Web-)Anwendung in Bezug auf eine Einwilligung in und einen Widerspruch gegen Datenverarbeitungen zum Zweck personalisierter Werbung enthält. Die Mitglieder von IAB Europe sind verpflichtet, der IAB Europe auf Anfrage alle Informationen zu übermitteln, die es IAB Europe ermöglichen, die Nutzer zu identifizieren, deren Daten Gegenstand eines TC-Strings sind. Wie bereits in seiner Entscheidung zu IP-Adressen, kommt es dem EuGH darauf an, ob die Stelle über rechtliche Mittel verfügt, die es ihr ermöglichen, eine bestimmte natürliche Person anhand der Informationen zu identifizieren, die ihr eine andere Stelle zur Verfügung stellt.[14]
- **Pressemitteilung des Europäischen Amtes für Betrugsbekämpfung (OLAF):** Der relative Maßstab des Personenbezugs verliert an Relevanz, wenn eine Stelle beabsichtigt, Daten zu **veröffentlichen**. So hat der EuGH zu einer Pressemitteilung des Europäischen Amtes für Betrugsbekämpfung (OLAF) entschieden, dass aus Sicht der veröffentlichenden Stelle auch dann personenbezogene Daten vorliegen können, wenn die Mitteilung keine direkten

[11] EuGH, Urt. v. 19.10.2016 – C-582/14, Rn. 43.
[12] EuGH, Urt. v. 19.10.2016 – C-582/14, Rn. 47ff.
[13] EuGH, Urt. v. 19.10.2016 – C-582/14, Rn. 45f.
[14] EuGH, Urt. v. 7.3.2024 – C-604/22, Rn. 48f.

Bezüge zu natürlichen Personen enthält. Dies nämlich dann, wenn die **Adressaten und Empfänger** solcher Mitteilungen über Mittel verfügen, die sie wahrscheinlich, d.h. mit verhältnismäßigem Zeit- und Kostenaufwand, zur Identifizierung der in solchen Mitteilungen nicht direkt in Bezug genommenen Personen nutzen werden.[15]

Zusammenfassend

> Daten haben einen Personenbezug und sind damit nicht anonym, wenn die Identifizierung einer Person wahrscheinlich ist, d.h. die verarbeitende Stelle
>
> - mit zumutbarem **Aufwand** an Zeit und Kosten
> - in rechtlich **zulässiger** Art und Weise
> - selbst oder mit **Hilfe** anderer Stellen
>
> eine natürliche Person identifizieren kann.

kommt es nach der Rechtsprechung des EuGH für die Frage des Personenbezugs darauf an, ob die verarbeitende Stelle eine natürliche Person wahrscheinlich, d.h. mit zumutbarem Aufwand an Zeit und Kosten und in rechtlich zulässiger Art und Weise selbst oder mit Hilfe anderer Stellen, identifiziert.

Wird dagegen eine Information für eine unbestimmte Vielzahl von Empfängern offengelegt, kommt es für die Frage des Personenbezugs dieser Information nicht allein auf die offenlegende Stelle, sondern auch auf die Empfänger und darauf an, ob diese die offengelegte Information wahrscheinlich mit zusätzlichen Informationen verknüpfen und die natürliche Person identifizieren werden.

> Wenn die Daten **veröffentlicht** oder anders **öffentlich zugänglich** sind, kommt es auch darauf an, ob die Empfänger die Person anhand der Daten wahrscheinlich identifizieren.

2.2 Pseudonymisierte Daten

Die DSGVO enthält keine Definition für pseudonyme Daten, aber für den Vorgang der Pseudonymisierung. Danach ist **Pseudonymisierung** die Verarbeitung personenbezogener Daten in einer Weise, dass diese ohne Hinzuziehung **zusätzlicher Informationen** nicht mehr einer

[15] EuGH, Urt. v. 7.3.2024 – C-479/22, Rn. 49, 55, 61.

spezifischen betroffenen Person zugeordnet werden können.[16] Eine Pseudonymisierung lässt den Personenbezug nicht entfallen,[17] da dieser mit Hilfe der genannten „zusätzlichen Informationen" noch hergestellt werden kann. Dabei handelt es sich im Allgemeinen um eine **Zuordnungsregel** oder **Zuordnungstabelle**, mit der sich die verwendeten **Pseudonyme** identifizierenden Angaben zu den Betroffenen wie Name und Geburtsdatum zuordnen lassen. Für eine wirksame Pseudonymisierung müssen die zusätzlichen Informationen gesondert aufbewahrt werden (siehe 2.2.1). Zudem müssen technische und organisatorische Maßnahmen ergriffen werden, die gewährleisten, dass diese Trennung nicht unberechtigt aufgehoben wird und die pseudonymisierten Daten nicht einer identifizierten oder identifizierbaren Person zugewiesen werden können (siehe 2.2.2). Bei der Bestimmung der erforderlichen technisch-organisatorischen Maßnahmen spielt das Konzept der **Pseudonymisierungs-Domäne** einer Rolle: dies ist der Bereich, in dem die Schutzwirkung der Pseudonymisierung greifen soll (siehe 2.2.3).[18]

2.2.1 Getrennte Aufbewahrung von personenbezogenen und identifizierenden Daten

Voraussetzung ist zunächst, dass die Daten über eine (unbestimmte) natürliche Person **von den Daten getrennt werden**, die den Rückschluss auf eine bestimmte Person zulassen. Daten über eine (unbestimmte) Person sind beispielsweise im Rahmen der Anmeldung für eine Fortbildungsveranstaltung Daten wie Anmeldezeitpunkt, Name und Preis der Fortbildungsveranstaltung. Daten, die den Rückschluss auf eine bestimmte Person zulassen, sind beispielsweise Name, Vorname und Anschrift. Für eine Pseudonymisierung müssen also die die Bestimmung der Person ermöglichenden Daten wie Name, Vornamen und Anschrift von den übrigen Anmeldedaten getrennt werden.

Um zu gewährleisten, dass diese Trennung nicht unumkehrbar ist, muss ein neues Datum, ein **Pseudonym**, generiert werden. Dieses ist über eine Zuordnungsregel oder eine Zuordnungstabelle mit den identifizierenden Daten verknüpft. Beispiele für Pseudonyme sind Kennungen wie Personal-, Steuer-, Matrikel- und Sozialversicherungsnummer, aber ggf. auch Vorgangs-, Rechnungs-, Antragsnummer oder ein Aktenzeichen. Das Pseudonym darf mit den (nicht

[16] Vgl. Art. 4 Nr. 5 DSGVO.
[17] Vgl. Erwägungsgrund 26 Satz 2 DSGVO.
[18] EDPB, Guidelines 01/2025 on Pseudonymisation, adopted on 16 January 2025, S. 4.

identifizierenden) Daten über eine Person verknüpft aufbewahrt werden und bildet mit diesen zusammen die **pseudonymisierten Daten**.
Dagegen muss das **Zusatzwissen**, d.h. die Verknüpfung des Pseudonyms mit den identifizierenden Daten von den pseudonymisierten Daten abgesondert und **getrennt aufbewahrt werden**. Wählt man im zuvor genannten Beispiel eine Antragsnummer, die mit Namen, Vornamen und Anschrift einer Person verknüpft ist, muss die Anmeldenummer als Pseudonym zweimal aufbewahrt werden: Einerseits mit den Anmeldedaten (Anmeldezeitpunkt, Name und Preis der Fortbildungsveranstaltung) und andererseits getrennt davon von mit den identifizierenden Daten (Namen, Vornamen und Anschrift).

2.2.2 Technische und organisatorische Maßnahmen

Eine Pseudonymisierung setzt nicht mehrere Beteiligte voraus, sondern kann auch **innerhalb einer datenschutzrechtlich verantwortlichen Stelle** vorgenommen werden.[19] Denn die Pseudonymisierung setzt nicht voraus, dass das Pseudonym und die mit ihm verknüpften identifizierenden Daten bei einer von der verantwortlichen Stelle organisatorisch getrennten Stelle aufbewahrt werden. Erforderlich ist lediglich, dass die datenschutzrechtlich verantwortliche Stelle **gesonderte Zugriffsberechtigungen** für diesen Datensatz vorsieht.[20] Weitergehende organisatorische und technische Maßnahmen zur Beschränkung des Zugriffs sind möglich und ggf. erforderlich.[21]

2.2.3 Pseudonymisierungs-Domäne

Diese richtet sich darauf, dass durch Trennung von Zuordnungsregel und pseudonymisierten Daten die Identifizierung der Betroffenen verhindert wird. Verantwortliche können einen Kontext definieren, in welchem diese Schutzwirkung erzielt wird, weil nur die pseudonymisierten Daten, nicht aber das Zusatzwissen verfügbar ist – die sog. **Pseudonymisierungs-Domäne**.[22]
Die Pseudonymisierungs-Domäne kann Stellen innerhalb des Verantwortlichen (z.B. bestimmte Abteilungen) umfassen oder auch außerhalb (z.B. Auftragsverarbeiter oder sonstige Dritte). Für

[19] Vgl. Erwägungsgrund 29 Satz 1 DSGVO.
[20] Vgl. Erwägungsgrund 29 Satz 2 DSGVO.
[21] Vgl. Art. 32 DSGVO.
[22] EDPB, Guidelines 01/2025 on Pseudonymisation, adopted on 16 January 2025, S. 4.

die Frage des Personenbezugs kommt es darauf an, ob die pseudonymisierten Daten und das Zusatzwissen kombiniert werden könnten unter Berücksichtigung aller Mittel, die von dem Verantwortlichen oder einer anderen Person wahrscheinlich genutzt werden.[23]

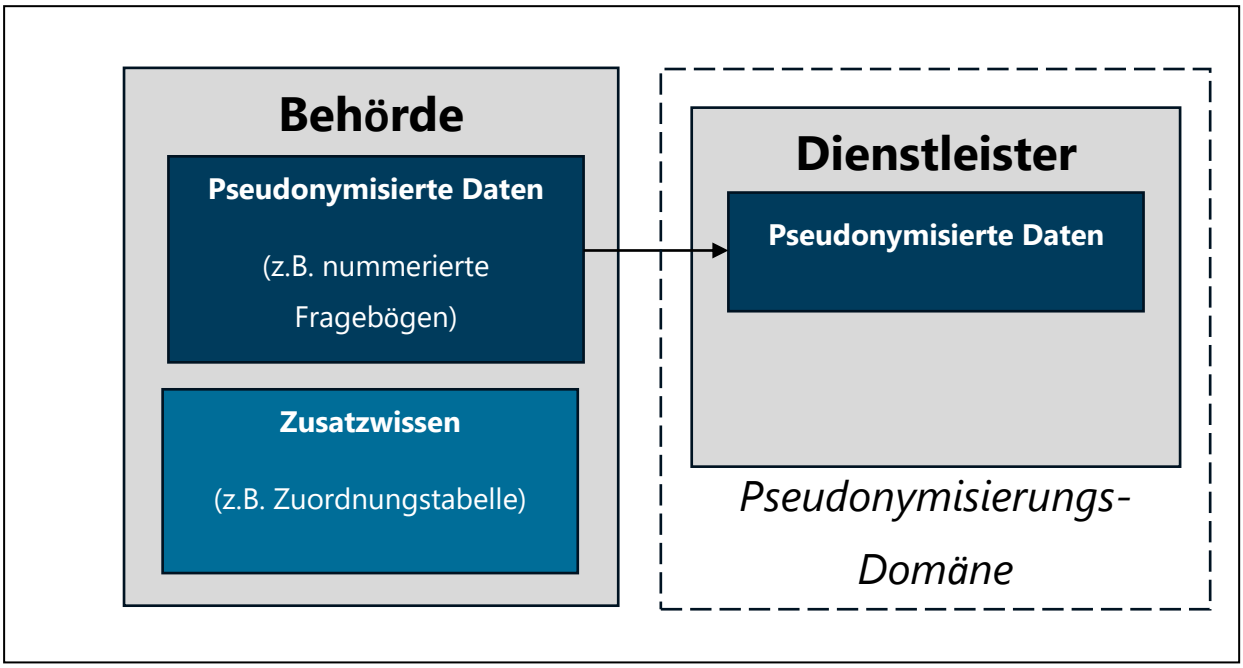

Abbildung 1: Pseudonymisierung einer Umfrage zur datensparsamen Auswertung durch einen externen Dienstleister

2.3 Anonyme und anonymisierte Daten

Die DSGVO erläutert den Begriff der anonymen Daten nur in ihren Erwägungsgründen.[24] Danach sind **anonyme Daten** Informationen, die sich nicht auf eine identifizierte oder identifizierbare natürliche Person beziehen.[25] Damit ist „anonyme Daten" der Gegenbegriff zu „personenbezogene Daten". Zu den anonymen Daten zählen **anonymisierte Daten**, d.h. personenbezogene Daten, die in einer Weise verarbeitet worden sind, dass die betroffene Person nicht oder nicht mehr identifiziert werden kann. Bei anonymisierten Daten bestand also ursprünglich ein Personenbezug, der im Nachgang entfernt worden ist. Der Prozess, in dem der

[23] EDPB, Guidelines 01/2025 on Pseudonymisation, adopted on 16 January 2025, S. 10 Rn. 22.
[24] Vgl. Erwägungsgrund 26 Satz 5 DSGVO.
[25] Vgl. Erwägungsgrund 26 Satz 5 DSGVO.

Personenbezug entfernt wird und der somit aus personenbezogenen Daten anonyme Daten macht, heißt **Anonymisierung**.[26]

Beispiele für anonyme und nicht anonymisierte Daten:

- Informationen, die sich allein auf Sachen beziehen, wie die Eigenschaften eines Produkts oder einer Dienstleistung.
- Daten, die sich auf Personengruppen (z.B. Vereine und Gesellschaften) beziehen und sich nicht zugleich aufgrund ihrer Art auf eine einzelne Person beziehen lassen.
- Statistische Daten, die von vornherein ohne Personenbezug erhoben werden (z.B. Induktionsspulen in der Straße zur Zählung der Fahrzeuge; Aufruf-Zähler auf Webseiten ohne weitere Datenerfassung).

Beispiele für anonyme und anonymisierte Daten:

- Daten über die Nutzung einer Website oder eines Online-Dienstes (z.B. Sitzungsdauer, Klickverläufe, Heat Maps, Fehlermeldungen), die zunächst in pseudonymen Nutzungsprofilen erhoben wurden (z.B. verknüpft mit einer eindeutigen User-ID oder einer ggf. personenbeziehbaren IP-Adresse), und die im Anschluss durch Entfernen oder Kürzen der möglichen Identifier anonymisiert wurden (z.B. durch Kürzen von IP-Adresse und Referrer-URL) und die im Übrigen keine personenbezogenen Inhaltsdaten enthalten (z.B. Name und Adresse oder Eingaben aus Freitextfeldern, bei denen mit personenbezogenen Eingaben gerechnet werden muss). Analysetools für Websites lassen sich in der Praxis regelmäßig so konfigurieren, dass mögliche personenbezogene Daten entsprechend anonymisiert (z.B. maskiert, gekürzt) werden.
- Antragsdaten, die lediglich noch das Antragsdatum und den Wohnort des Antragstellers enthalten und aus denen weitere Identifier entfernt wurden (z.B. Name und Postanschrift) und bei denen keine Verknüpfung zu anderen Tabellen/Datenbanken wahrscheinlich ist (zum Beispiel keine Verknüpfung zu anderen Daten über Pseudonyme wie Personalnummer

[26] Vgl. Art. 2 Nr. 8 Open Data und PSI-Richtlinie (RL (EU) 2019/1024).

oder Antragsnummer und auch keine Verknüpfung über Daten etwa aus einem Data Warehouse (DWH)).

Voraussetzungen der Anonymisierung: Die DSGVO lässt offen, unter welchen Voraussetzungen davon ausgegangen werden, dass Daten sich nicht oder nicht mehr auf eine identifizierte oder identifizierbare Person beziehen.[27] Diese Lücke kann durch Rückgriff auf die Kriterien zur Bestimmung des Personenbezugs von Daten geschlossen werden.[28] Unter Berücksichtigung der Rechtsprechung des EuGH zu IP-Adressen ist somit grundsätzlich[29] dann von anonymen Daten auszugehen, wenn die Stelle, die die Daten verarbeitet, nur mit unzumutbarem Aufwand an Zeit oder Kosten oder nur mit illegitimen Mitteln in der Lage ist, die Identifizierung der natürlichen Person selbst oder mit Hilfe anderer Stellen vorzunehmen. Auch für die Anonymität von Daten kommt es auf die Technologie an, die zum Zeitpunkt der Verarbeitung dieser Daten verfügbar ist. Eine einmal herbeigeführte Anonymität bleibt somit nicht zwangsläufig bestehen und muss regelmäßig überprüft werden.

2.4 Datenbanken

Datenbanken: Eine Datenbank ist eine strukturierte Sammlung von Daten, die organisiert, gespeichert und verwaltet wird, um auf die Daten effizient zugreifen, diese abfragen und sie aktualisieren zu können. Datenbanken werden in vielen verschiedenen Anwendungen eingesetzt, einschließlich KI-Systemen.

Eine Datenbank besteht aus zwei Komponenten: einer Verwaltungssoftware, genannt **Datenbank-Management-System (DBMS)**, und der **Datenbank (DB)** im engeren Sinn. Das DBMS organisiert intern die strukturierte Speicherung der Daten und ermöglicht es Benutzenden, Daten in der Datenbank hinzuzufügen, zu ändern oder zu löschen. Es legt fest, welche Benutzenden inwieweit auf eine Datenbank zugreifen können, und ermöglicht, Abfragen (sog. Querys) auf einer Datenbank vorzunehmen (z.B. Abfrage zur Anzeige der Anzahl der Anträge in einem Onlinedienst im letzten Halbjahr).

[27] Vgl. Erwägungsgrund 26 Satz 5 DSGVO.
[28] Vgl. Erwägungsgrund 26 Satz 4 DSGVO.
[29] Liegt ein Fall der Offenlegung gegenüber einer unbestimmten Vielzahl von Empfängern vor, ist auch maßgeblich, ob die Empfänger die offengelegte Information wahrscheinlich mit zusätzlichen Informationen zur Identifizierung einer natürlichen Person verknüpfen. Siehe oben Ziff. II.1.

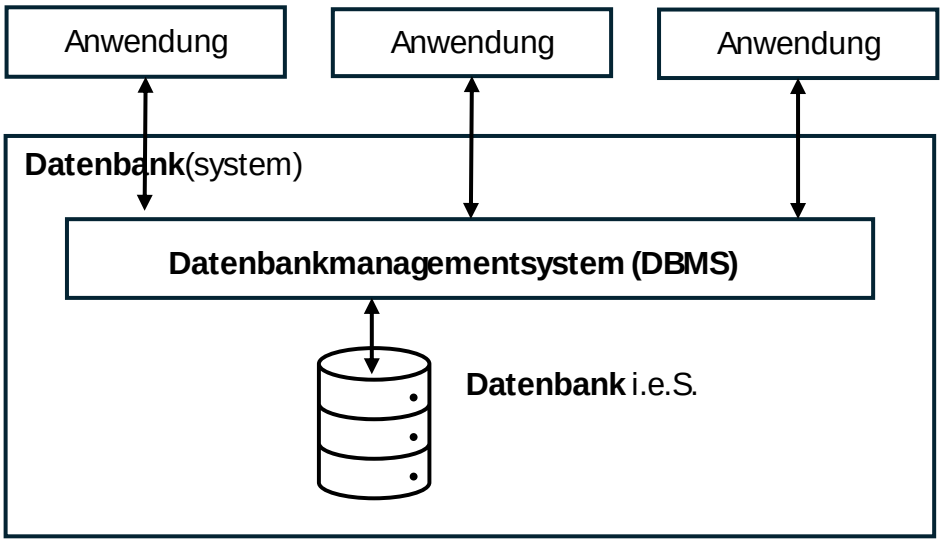

Abbildung 2: : Aufbau eines Datenbanksystems

Die in einer Datenbank abgelegten Daten werden vor der Speicherung in einem vorgegebenen Format strukturiert. Es handelt sich damit um **strukturierte Daten** (im Gegensatz zu unstrukturierten Daten, die mehr oder weniger zufällig in größerer Menge vorhanden sind, zum Beispiel eine nicht näher geordnete Ansammlung von Bild- oder sonstigen Dateien auf einem Datenträger). Die konkrete Struktur einer Datenbank wird durch das verwendete **Datenbankmodell** bestimmt. Die gebräuchlichsten Formen sind **relationale Datenbanken**, die auf Tabellen basieren. In den Tabellen (auch Relationen genannt) sind die Daten in einem Format enthalten, das durch das Relationsschema (entsprechend der Kopfzeile der Tabelle) vorgegeben wird. Die in den Zeilen enthaltenen Datensätze (auch *Tupel*) besitzen ein oder mehrere Schlüsselattribute mit eindeutigen Werten und darüber hinaus eine beliebige Menge von weiteren Attributen.

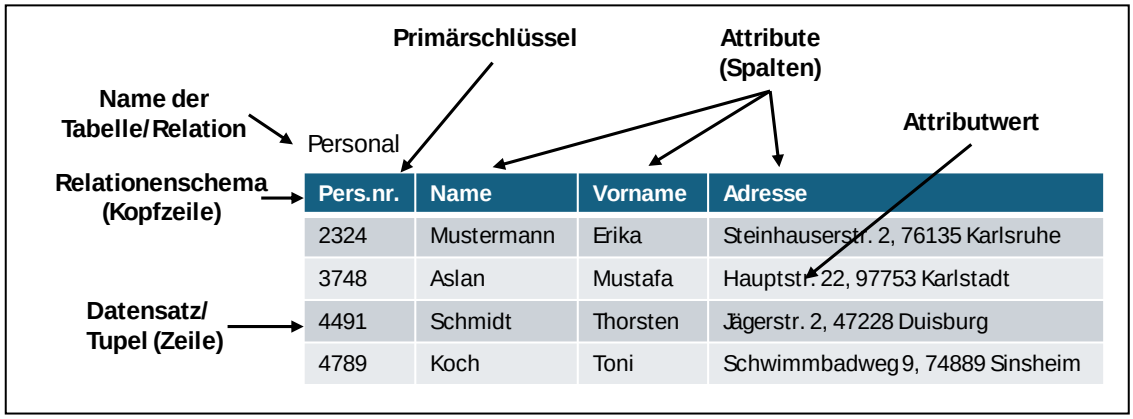

Abbildung 3: Tabelle in einer relationalen Datenbank

2.5 KI-Systeme

Die nachfolgenden Grundbegriffe aus dem Bereich der KI berücksichtigen die Terminologie der europäischen KI-Verordnung (KIVO), welche zwischen KI-Modell und KI-System unterscheidet.

Künstliche Intelligenz (KI): Die KI ist das Teilgebiet der Informatik, welches sich mit der Automatisierung intelligenten Verhaltens und dem maschinellen Lernen befasst.

KI-Algorithmus: Ein Algorithmus ist eine automatisierte Anweisung – ein „Rezept" zur Lösung einer Aufgabe durch Rechenschritte. Die KI-Forschung hat diverse Methoden entwickelt, um Aufgabenstellungen aus den Bereichen Suchen, Planen, Optimieren, Logisches Schlussfolgern und Abstrahieren algorithmisch zu lösen. Die bekannten KI-Algorithmen lassen sich einteilen in symbolische und neuronale KI (ein prominenter Unterfall neuronaler KI sind Large-Language-Models (LLM)):

- **Symbolische KI** folgt einem „Top-Down-Ansatz". Ausgangspunkt ist das für die Problemstellung relevante Wissen. Dieses wird in Form von abstrakten Symbolen bzw. Begriffen repräsentiert und dann mithilfe von logischen Regeln verarbeitet. Die Funktionsweise einer solcher KI lässt sich grundsätzlich gut erklären, da sie auf einer begrifflichen Ebene operiert und die Anwendung der Regeln für Menschen nachvollziehbar ist. Zu den symbolischen KI-Algorithmen zählen z.B. Entscheidungsbäume und andere mathematisch-statistische Verfahren zur Optimierung und Klassifizierung. Auch sog. genetische Algorithmen, die an die biologische Evolution angelehnt sind, fallen in diese Kategorie.

- **Neuronale KI** folgt einem „Bottom-Up-Ansatz". Inspiriert vom menschlichen Gehirn werden Teilaspekte davon im Rechner nachgebildet, um Intelligenzleistungen zu erzielen. Insbesondere werden „künstliche Neuronen" simuliert und zu neuronalen Netzen (NN) verbunden. Jedes Neuron entspricht dabei einer einfachen numerischen Funktion, mit der aus einer Menge von Eingabewerten (die von den vorgeschalteten Neuronen zugeliefert werden) ein Anregungszustand berechnet und an die nachgeschalteten Neuronen übermittelt wird. Auf diese Weise fließen Informationen durch das NN von einer „Eingabeschicht" bis zu den äußersten Neuronen, die als „Ausgabeschicht" interpretiert werden und die „Antwort" des NN auf eine Eingabe liefern. Trotz dieses einfachen Grundprinzips sind NN zu erstaunlichen Leistungen fähig von der Mustererkennung bis hin zur Generierung von Text, Sprache, Bildern etc.

  Ein NN wird durch Anzahl, Anordnung und Vernetzung der Neuronen charakterisiert (Topologie). Der konkrete Zustand wird durch weitere Parameter beschrieben, insbesondere durch die Stärke der jeweiligen Neuronenverbindungen (Gewichtung) und durch die Grundanregung der einzelnen Neuronen (Bias).

  Im Gegensatz zur symbolischen KI liegt ein Nachteil sämtlicher NN in der prinzipiell **schlechten Erklärbarkeit** ihrer Ergebnisse für den Menschen. Der innere Zustand eines NN lässt sich zunächst nur auf der Ebene der Gewichte und Bias-Werte beobachten. Diese numerischen Werte lassen sich im Allgemeinen nicht direkt in für Menschen verständliche Begriffe und Symbole übersetzen. Somit ist es gerade bei komplexen NN (wie LLM, dazu sogleich) schwer bis unmöglich, die konkreten Ausgaben in für Menschen nachvollziehbarer Weise durch eine Anwendung von Regeln und logischen Schlüssen herzuleiten.
- **Large-Language-Models (LLM)**: Diese Form von NN ist aufgrund von Durchbrüchen in der Forschung in jüngerer Zeit in den Fokus gerückt. LLM sind große NN (mit Milliarden oder Billionen von Parametern), die unter hohem Recheneinsatz auf großen Mengen von Texten trainiert werden und daraus statistische Beziehungen zwischen Begriffen lernen. Sie können im Anschluss Text-Eingaben (Prompts) durch Finden des jeweils wahrscheinlichsten nächsten Wortes beliebig schrittweise fortschreiben und damit verschiedenste Anforderungen im Bereich der Generierung von Texten erfüllen. Aufgrund ihrer Vielseitigkeit fallen LLM

nach der KI-Verordnung unter den Begriff der „KI-Modelle mit allgemeinem Verwendungszweck".[30]

KI-Modell:[31] Ein KI-Modell ist ein Computerprogramm, das auf Basis von KI-Algorithmen menschliche Fähigkeiten wie logisches Denken, Lernen, Planen und Kreativität imitieren kann. Insbesondere gibt es KI-Modelle, die nach einer Trainingsphase in der Lage sind, Muster in Datensätzen zu erkennen oder auf Basis von Eingabedaten Entscheidungen zu treffen. KI-Modelle sind in der Regel in KI-Systeme integriert und bilden deren Kernkomponente.

KI-System: KI-Systeme sind Software- und Hardwaresysteme, die KI nutzen, um aus erhaltenen Eingaben „rational" Ausgaben abzuleiten, die je nach Einsatzzweck z.B. als Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen interpretiert und in der physischen oder virtuellen Welt verwendet werden können.[32] Den Kern eines KI-Systems bildet ein KI-Modell, das jedoch um weitere Systembestandteile – wie z.B. eine Nutzerschnittstelle oder eine zusätzliche Wissensdatenbank – ergänzt wird. KI-Systeme können mehr oder weniger autonom sein.[33]

[30] Vgl. Art. 3 Nr. 63 KIVO.
[31] Vgl. Erwägungsgrund 97 KIVO.
[32] Vgl. Art. 3 Abs. 1 KIVO.
[33] Vgl. Art. 3 Abs. 1 KIVO.

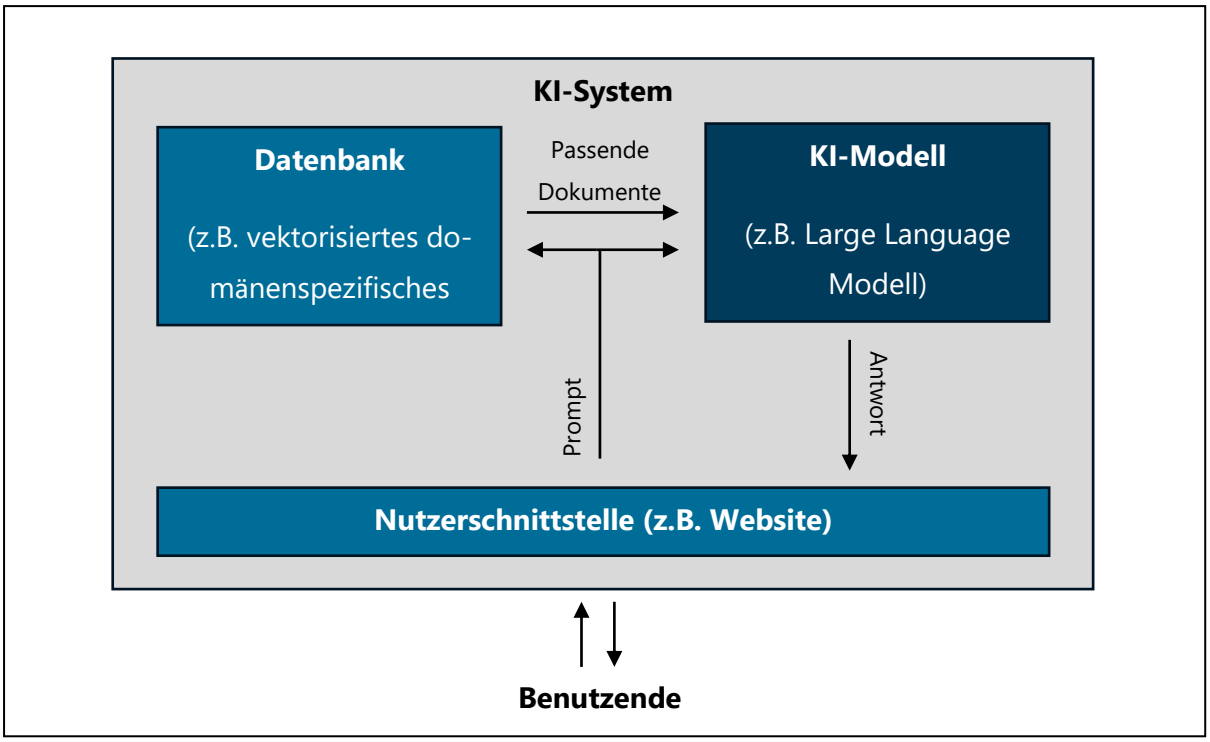

Abbildung 4: Beispiel eines KI-Systems, das ein KI-Modell und weitere Teilsysteme umfasst und nach dem RAG-Prinzip („Retrieval Augmented Generation") arbeitet

Maschinelles Lernen (ML) und **Training**: ML ist ein Teilgebiet der KI und befasst sich damit, wie Programme Wissen aus Daten lernen können, um praktische Probleme zu lösen. Einen konkreten Lernprozess nennt man Training. Dabei bildet ein Lernalgorithmus vorgegebene Beispieldaten auf ein mathematisches Modell ab. Während des Trainings wird das Modell durch den Lernalgorithmus so angepasst (z.B. durch Anpassung der Parameter eines neuronalen Netzes), dass die Ergebnisse des Modells die gegebene Aufgabe möglichst gut lösen. Nach dem Training ist der gefundene Lösungsweg im Modell gespeichert. Dieser Lösungsweg ist nicht vom Menschen explizit programmiert worden, sondern wurde durch den Lernalgorithmus aus den Trainingsdaten abgeleitet. Beim maschinellen Lernen werden drei Vorgehensweisen unterschieden:

- **Überwachtes Lernen:** Beim überwachten Lernen sind die Trainingsdatensätze „gelabelt", d.h. sie enthalten explizit die Information über die zu erlernende Kategorie. Soll etwa eine KI lernen, Bilder von Hunden und Katzen zu unterscheiden, so würde man beim überwachten Lernen mit Bildern arbeiten, die jeweils mit „Hund" oder „Katze" markiert sind. Der

Lernalgorithmus passt dann die Modell-Parameter an, damit die Ausgabe des Modells möglichst mit den Labeln übereinstimmt.

- **Unüberwachtes Lernen:** Beim unüberwachten Lernen erstellt der Lernalgorithmus aus den Eingabedaten ein Modell, welches Zusammenhänge und erkannte Kategorien enthält (z.B. erkannte Cluster/Häufungen in den Daten). Es wird keine Klassifikation (wie im Beispiel zuvor „Hund" und „Katze") vorgegeben, sondern ggf. vom Algorithmus entwickelt.
- **Bestärkendes Lernen:** Das bestärkende Lernen wird eingesetzt für Lernsysteme, die als „Agenten" gestaltet sind, d.h. die eine Umgebung wahrnehmen und darin Aktionen ausführen können. Für gewünschte Aktionen erhält das Lernsystem eine Belohnung. Es passt sich daraufhin an und versucht, eine Strategie zu entwickeln, um möglichst viele Belohnungen zu erhalten. Bestärkendes Lernen wird z.B. beim Lernen von Spielen oder bei der Entwicklung von Systemen für autonomes Fahren eingesetzt.

2.6 Datenkategorien bei Entwicklung und Nutzung von KI

Bei der Entwicklung und Nutzung von KI, die auf „Machine Learning" beruht, sind die folgenden Datenkategorien zu unterscheiden.

2.6.1 Entwicklung

Trainingsdaten:[34] Diese Daten werden in der Trainingsphase eingegeben. Sie sind so aufgebaut wie die Daten, die im späteren Betrieb in das KI-System eingegeben werden sollen. Das KI-System lernt daraus bestimmte Zusammenhänge und Regelmäßigkeiten und speichert diese, indem es seine Modellparameter anpasst („Machine Learning"). Beim überwachten Lernen sind die Trainingsdaten bereits nach den zu erlernenden Kategorien gelabelt. Hingegen funktioniert das unüberwachte Lernen mit ungelabelten Trainingsdaten, in denen das KI-System selbständig Strukturen und Cluster erkennt.

Validierungsdaten:[35] Validierungsdaten sind genauso aufgebaut wie Trainingsdaten. Sie werden nach dem Training eingegeben, um die Erkennungsgenauigkeit des KI-Systems zu messen (Validierung) und dessen Einsatztauglichkeit zu bestimmen. Durch die Trennung von Trainings-

[34] Vgl. Art. 3 Nr. 29 KIVO.
[35] Vgl. Art. 3 Nr. 30 KIVO.

und Validierungsdaten wird verhindert, dass das KI-System allein auf die Trainingsdaten optimiert wird (sog. Overfitting).

Wissensbasis des KI-Systems: Das durch das Training generierte Wissen wird im KI-System gespeichert. Bei neuronalen Netzen wie etwa Large-Language-Models (LLM) geschieht dies in Form von Modell-Parametern (der Netztopologie, den Gewichtungen von Neuronenverbindungen und Bias-Werten). In der Regel ist das so gespeicherte Wissen abstrakt und enthält keine Informationen über identifizierbare Personen, auch wenn personenbezogene Trainingsdaten genutzt wurden. Unter Umständen können aber auch natürliche Personen in dem Wissen des KI-Systems repräsentiert sein (so enthält etwa ChatGPT erkennbar Wissen über prominente Personen).

2.6.2 Nutzung

Eingabedaten:[36] Die Eingabedaten umfassen alle Informationen, die dem KI-System während des Betriebs präsentiert werden. Dies können etwa Texte (z.B. E-Mails, Dokumente), Bilder, Videos oder Tonaufnahmen sein. Je nach Einsatzzweck können die Eingabedaten über Sensoren erfasst (z.B. Smart Cams, Mikrofone), aus Dateien oder von IT-Schnittstellen gelesen oder vom Nutzer eingegeben werden. Bei LLM, die mit dem Nutzer in Dialogform interagieren (z.B. ChatGPT oder Copilot), nennten man die Eingabedaten auch „Prompt".

Ausgabedaten: Diese werden von der KI ausgegeben bzw. erzeugt (Output). Bei KI, die zur Klassifizierung eingesetzt wird, ist dies die vom System vorgenommene Zuordnung der Eingabedaten zu den Klassen. Bei der sog. generativen KI ist der Output komplexer und kann u.a. in Texten, Sprache, Bildern, Musik oder Videos bestehen.

[36] Vgl. Art. 3 Nr. 33 KIVO.

3 Anonymisierungsbedarf bei KI und gängige (KI-)Verfahren zur Anonymisierung

3.1 Bedarf zur Anonymisierung der einzelnen Datenkategorien

In der Praxis kann bei allen eben genannten Datenkategorien ein Bedarf nach Anonymisierung bestehen:

Trainingsdaten: Bei personenbezogenen Trainingsdaten kann eine Anonymisierung erforderlich sein, um das KI-Training als Verarbeitungsvorgang zu rechtfertigen (Ermöglichungsfunktion, s.o.). Außerdem erfordert mit Blick auf die Schutzfunktion der Anonymisierung der Grundsatz der Datensparsamkeit, anonyme Trainingsdaten zu verwenden, wenn dies möglich ist (d.h. insbesondere die Trainingsziele auch mit anonymen Daten erreicht werden können) und zumutbar ist (d.h. für die Anonymisierung kein Aufwand erforderlich ist, der außer Verhältnis steht zu dem Kosten- und Zeitersparnis aus der KI-Nutzung).

Validierungsdaten: Hier gilt dasselbe wie für die Trainingsdaten. Die Entscheidung, ob zu anonymisieren ist, kann jedoch anders ausfallen, z.B. wenn mit anonymen Validierungsdaten das KI-System nicht hinreichend validiert werden kann.

Wissensbasis des KI-Systems (bzw. KI-Modells): KI-Modelle gewinnen durch das Training abstraktes „Wissen", welches sie zur Lösung ihrer Aufgabenstellung befähigt. Dieses Wissen ist jedoch in einer Weise gespeichert, die sich grundlegend etwa von einer geordneten Datenbank unterscheidet. Vielmehr ist das Wissen innerhalb der KI-Modelle typischerweise in einer Form kodiert, die es schwer bis unmöglich macht, zu lokalisieren, wo und wie bestimmte Informationen gespeichert sind. So wird der Modellzustand neuronaler Netze durch eine Reihe mathematischer Werte beschrieben (Neuronengewichte, Biasgewichte etc.). Der Aufwand, um nachzuvollziehen, welche Zustände mit welchem (Fakten-)Wissen korrelieren, steigt mit der Komplexität der Modelle stark an. Zudem werden einzelne Fakten in komplexen Modellen nicht durch einzelne Werte bzw. Neuronen gespeichert, sondern durch eine Überlagerung von Zuständen mehrere Neuronen (Superposition). Schließlich kommt hinzu, dass die KI-Modelle grundsätzlich keine spezifischen/exakten Fakten lernen, sondern ihr Wissen durch **Abstraktion aus den Trainingsdaten** gewinnen und in stochastischer/komprimierter Weise abspeichern. Der Modellzustand repräsentiert insofern **unscharfes Wissen**. Es werden

Wahrscheinlichkeitswerte generiert.[37] Dies erschwert es erheblich, die konkrete Funktionsweise von KI-Modellen im Einzelfall nachvollziehbar zu machen. So ist es nahezu unmöglich, im Einzelnen darzulegen, welche Kette von Schlüssen innerhalb eines LLM zu einer bestimmten Antwort geführt haben.

Aufgrund dieser Eigenschaften von KI-Modellen ist umstritten, ob sie überhaupt personenbezogenen Daten enthalten können.[38] Nach dem Europäischen Datenschutzausschuss sind jedenfalls KI-Modelle mit einer datenschutzrechtlich anonymen Arbeitsweise denkbar.[39] Allerdings wird eine Anonymisierung innerhalb des KI-Modells regelmäßig technisch unmöglich oder sinnlos sein. Dies gilt aber nicht notwendig für das KI-System als Ganzes. So ist unstreitig, dass sowohl die Eingabe als auch die Ausgabe von KI-Systemen (einschließlich von KI-Systemen, die auf einem LLM beruhen) personenbezogene Daten enthalten kann. Zudem können innerhalb eines KI-Systems neben dem Modell weitere Subsysteme mit personenbezogenen Daten (z.B. eine Datenbank, die vom Modell „befragt" werden kann) enthalten sein. Das Vorhandensein personenbezogenen Wissens muss bei der Beurteilung des Systems berücksichtigt werden (insbesondere, wenn ein weiteres Training stattfindet). Sofern personenbezogene Daten in der Wissensbasis bzw. im KI-Modell enthalten sind, ist es ggf. möglich, durch geeignete Maßnahmen dafür zu sorgen, dass diese Daten nicht in die Ausgabedaten gelangen (z.B. durch technische Eingangs-/Ausgangsfilter – vgl. 5.1.2 zur „Differential Privacy" oder durch organisatorische Anweisungen). Dadurch lässt sich unter den konkreten Umständen der Nutzung für die Wissensbasis im KI-Modell aus Sicht einer benutzenden Stelle Anonymität gewährleisten.

Eingabedaten: Die Eingabe von personenbezogenen Daten in ein KI-System führt zu verschiedenen Verarbeitungsvorgängen (Verarbeitung zur Erzeugung der Ausgabe, ggf. Nachtraining zur Verbesserung des KI-Systems), die aus Datenschutzsicht zu prüfen sind. Dabei kann sich je nach den Umständen aus den Grundsätzen der Rechtmäßigkeit, Erforderlichkeit und Datensparsamkeit die Anforderung ergeben, die Daten vor der Eingabe zu anonymisieren. Eine Anonymisierung ist regelmäßig auch dann erforderlich, wenn das KI-System nicht unter der

[37] *EDSA*, Opinion 28/2024, 17. Dezember 2024, Rn. 38.
[38] Nach *Fuchs*, KIR 2024, 79, 80 weisen die im (LLM-)Modell gespeicherten Parameter keinen Personenbezug auf. Nach *Hansen/Walczak*, KIR 2024, 82, enthalten die in einem LLM repräsentierten Informationen in vielen Fällen personenbezogene Daten aus den Trainingsdaten.
[39] *EDSA*, Opinion 28/2024, 17. Dezember 2024.

Kontrolle des Verantwortlichen steht (z.B. über das Internet zugängliche generative KI ohne Auftragsverarbeitung).

Ausgabedaten: Wenn Ausgabedaten weiterverarbeitet – insbesondere an Dritte weitergegeben oder veröffentlicht – werden sollen, ist zu prüfen, ob sie personenbezogene Informationen enthalten, die zuvor zu anonymisieren sind.

3.2 KI zur Anonymisierung

Anonymisierung ist eine zentrale Maßnahme des Datenschutzes. Der Bedarf an der Anonymisierung von Datenbeständen ist daher nicht nur bei der Entwicklung und dem Einsatz von KI-Systemen, sondern auch bei anderen Verarbeitungsvorgängen hoch.

Da Anonymisierung je nach der Art der Daten einen erheblichen Aufwand bedeuten kann, gibt es verschiedene Ansätze zur Automatisierung. Dabei wird auch KI eingesetzt. Unter anderem sind dabei folgende Verfahren relevant:

Synthetische Daten: Bei dieser Methode werden anstelle von Echtdaten künstlich erzeugte Datensätze generiert und als Trainingsdaten verwendet. Synthetische Daten weisen grundsätzlich keinen Bezug zu realen Personen auf. Sie können mithilfe von generativer KI erstellt werden, die zunächst die erforderlichen Merkmale (Strukturen, statistische Verteilungen), aus den Echtdaten lernt und dann entsprechende synthetische Daten erzeugt. Die Herausforderung besteht darin, synthetisch Daten zu schaffen, die alle für das Training erforderlichen Merkmale von Echtdaten aufweisen.

Maskierung in Texten / Named Entity Recognition (NER): Eine Methode zur Anonymisierung von Texten ist die Maskierung von Identifiern (z.B. Namen, Anschriften, Geburtstagen, Ortsangaben, sonstige stark individualisierende Merkmale). Um sie maskieren zu können, müssen die Identifier zunächst erkannt werden. In der Computerlinguistik heißt diese Aufgabe „Named Entity Recognition". Es gibt gute frei verfügbare KI-Modelle für NER. Je nach Einsatzzweck müssen diese jedoch ggf. nachtrainiert werden, um im spezifischen Fall adäquate Ergebnisse zu liefern (z.B. bei besonderen Sprachen, speziellen Identifiern oder besonderen Kontexten).

Maskierung in Bildmaterial / Smart Cams: Eine häufige Anforderung bei Videoüberwachung ist, bestimmte Bereiche oder Personen im angezeigten Bild zum Schutz der Privatsphäre zu maskieren. Hierbei kommen KI-Algorithmen zum Einsatz, die u.a. sich bewegende Objekte oder Personen erkennen und überblenden können. Andere KI-basierte Verfahren berechnen aus den

Aufnahmen ein 3D-Modell der Realität (Digital Twin) und zeigen dieses mit einer simplifizierten Grafik an, so dass Personen nicht identifizierbar sind (Dummies, Strichmännchen).

4 Datenschutzrechtliche Rahmenbedingungen der Anonymisierung

4.1 Anonymisierung als Datenverarbeitung

Nach überwiegender Ansicht ist der Vorgang der Anonymisierung als solcher eine **Verarbeitung personenbezogener Daten**.[40] Dafür spricht, dass der Verarbeitungsbegriff in der DSGVO[41] weit definiert ist und jeden Vorgang im Zusammenhang mit personenbezogenen Daten erfasst. Die Folge dieser Ansicht ist, dass die DSGVO auf die Anonymisierung Anwendung findet und alle Datenschutzgrundsätze zu beachten sind. Insbesondere bedarf es einer **datenschutzrechtlichen Rechtfertigung**[42], um eine Anonymisierung durchführen zu können.

4.2 Rechtsgrundlage für die Anonymisierung

Als mögliche **Rechtsgrundlage für eine Anonymisierung** von Daten durch eine behördliche Einrichtung kommen – bei Vorliegen der jeweiligen Voraussetzungen – u.a. die Wahrnehmung öffentlicher Aufgaben[43], eine rechtliche Verpflichtung[44], eine Einwilligung[45] oder die Anbahnung oder Erfüllung eines Vertrages[46] in Betracht. Hingegen steht die Rechtsgrundlage des berechtigten Interesses[47], die im nicht-öffentlichen Bereich eine große praktische Bedeutung hat, für die Anonymisierung durch öffentliche Stellen nicht zur Verfügung.

[40] *Hornung/Wagner*, Anonymisierung als datenschutzrelevante Verarbeitung?, ZD 2020, 223, 224.
[41] Vgl. Art. 4 Nr. 2 DSGVO.
[42] Vgl. Art. 6 DSGVO.
[43] Vgl. Art. 6 Abs. 1 Buchstabe e DSGVO i.V.m. der landesdatenschutzrechtlichen Generalklausel (siehe unter anderem § 4 HmbDSG, § 3 NDSG, § 4 LDSG).
[44] Vgl. Art. 6 Abs. 1 Buchstabe c DSGVO.
[45] Vgl. Art. 6 Abs. 1 Buchstabe a DSGVO.
[46] Vgl. Art. 6 Abs. 1 Buchstabe b DSGVO.
[47] Vgl. Art. 6 Abs. 1 Buchstabe f DSGVO.

Für den Fall, dass eine **KI trainiert** werden soll, die anschließend durch Behörden zur Wahrnehmung der öffentlich-rechtlichen Verwaltungstätigkeit eingesetzt werden soll, gibt es in manchen Bundesländern inzwischen eine Rechtsgrundlage für die Verarbeitung personenbezogener Daten zu diesem Zweck.[48] Zum Teil wurde explizit eine Rechtsgrundlage für die Anonymisierung geschaffen.[49]

Hilfestellung bei der Schaffung neuer Rechtsgrundlagen für die Datenverarbeitung bietet der Rechtsgrundlagen-Generator des Kompetenzteams Datenschutz im Schwerpunktthema Datennutzung des IT-Planungsrates.

4.3 Anonymisierung als Zweckänderung

Wenn bereits bei einer Behörde vorhandene Daten – z.B. aus einem Fachverfahren –anonymisiert werden, um sie einem anderen Zweck zuzuführen – wie dem Training einer neuen KI – handelt es sich datenschutzrechtlich um einen Fall der **Zweckänderung**. Hierfür sind grundsätzlich weitere Voraussetzungen zu beachten[50].

Die rechtliche Systematik der Zweckänderung ist recht komplex und im Detail umstritten. Grundsätzlich gilt, dass eine Zweckänderung auf eine Einwilligung oder eine gesetzliche Erlaubnis gestützt werden kann.[51] Ist das nicht der Fall, kann die Zweckabweichung nach der DSGVO zulässig sein, wenn der neue Zweck **mit dem bisherigen Zweck kompatibel** ist. Im Rahmen dieser Kompatibilitätsprüfung wird die Verbindung zwischen den Zwecken, der Zusammenhang zwischen Erhebung und Weiterverarbeitung, die Art der Daten und Verarbeitungsfolgen sowie Garantien (u.a. Verschlüsselung und Pseudonymisierung) betrachtet.[52]

Umstritten ist, ob das Vorliegen dieser Voraussetzungen ausreicht, um die Zweckänderung zu rechtfertigen,[53] oder ob **zusätzlich eine Rechtsgrundlage** erforderlich ist (diese Auslegung

[48] Vgl. § 13 Abs. 2 HmbVwDiG, § 8 Abs. 1 und 2 ITEG-SH.
[49] Vgl. § 13 Abs. 2 HmbVwDiG, wonach auch die Anonymisierung besonderer Kategorien personenbezogener Daten nach Art. 9 Abs. 1 DSGVO gestattet ist.
[50] Vgl. Art. 6 Abs. 4 DSGVO. In den Landesdatenschutzgesetzen werden Fälle einer zulässigen Zweckabweichung geregelt. Vgl. hierzu etwa § 6 Abs. 2 HmbDSG, § 5 Abs. 1 und 2 LDSG BW, § 6 Abs. 1 BbgDSG.
[51] Eine zulässige Zweckänderung ist beispielsweise in Hamburg in § 13 Abs. 3 HmbVwDiG geregelt.
[52] Vgl. Art. 6 Abs. 4 DSGVO.
[53] Dafür spricht Erwägungsgrund 50 Abs. 2 DSGVO. Vgl. u.a. *Schulz*, in: Gola/Heckmann, Datenschutz-Grundverordnung – Bundesdatenschutzgesetz, 3. Auflage 2022, Art. 6 Rn. 133; *Schantz,* NJW 2016, 1841, 1844.

vertreten mehrheitlich die Aufsichtsbehörden). Die Frage bedarf aber dann keiner Entscheidung, wenn die Zweckänderung auf eine eigene gesetzliche Grundlage gestützt werden kann.[54]

5 Technische Rahmenbedingungen der Anonymisierung

5.1 Anonymitätsmaße

Aus Sicht des Datenschutzrechts ist der Personenbezug von Daten als Schwarz-/Weiß-Kriterium konzipiert, also binär. Liegt ein Personenbezug vor, so handelt es sich um personenbezogene Daten, auf die das Datenschutzrecht anwendbar ist, anderenfalls handelt es sich um anonyme Daten, auf die das Datenschutzrecht keine Anwendung findet. Personenbezug und Anonymität von Daten bilden also aus rechtlicher Perspektive ein Gegensatzpaar, bei dem es keine Zwischenzustände gibt (pseudonyme Daten sind in der rechtlichen Betrachtung personenbezogen).

In der wissenschaftlichen Forschung zum Datenschutz (Information Privacy Research) wird Anonymität hingegen im Allgemeinen als eine Größe verstanden, die für gegebene Daten den Aufwand zur Identifizierung einer Person angibt und die innerhalb eines Kontinuums liegt. Man kann diese Größe dementsprechend durch eine Zahl zwischen 0 (= keine Möglichkeit, einen Personenbezug herzustellen) und 1 (= offenkundiger Personenbezug) ausdrücken. Eine rechtliche Vorgabe, ab welchem Schwellwert von einem Personenbezug im Sinne der DSGVO auszugehen ist, gibt es zwar nicht. Gleichwohl ist es hilfreich, den „Grad der Anonymität" bzw. den Aufwand zur Herstellung eines Personenbezugs zu formalisieren und technisch messbar zu machen. Mit einem solchen Maß kann die Wirksamkeit von Anonymisierungstechniken als Maßnahmen zur Verringerung von Risiken für die Betroffenen beurteilt und eine objektive Grundlage für deren rechtliche Beurteilung geschaffen werden. Die Wissenschaft hat dazu verschiedene **Anonymitätsmaße** entwickelt.

5.1.1 k-Anonymität, l-Diversität und t-Geschlossenheit

Die Forschung zur Anonymität von Daten betrachtet eine Menge gleichartiger Datensätze (im Folgenden „Datenbank" oder „Tabelle") und stellt sich die Frage, welches Wissen daraus über

[54] In Hamburg ist mit § 13 Abs. 3 HmbVwDiG eine solche Grundlage geschaffen worden.

einzelne Individuen abgeleitet werden kann. Dabei werden die denkbaren Methoden („Angriffe“) berücksichtigt, mittels derer – ggf. auch unter Rückgriff auf Hintergrundwissen – Informationen aus der Tabelle gewonnen werden können. Drei grundlegende Ansätze, um die Erfolgsaussichten solcher Angriffe und damit die Anonymität einer Menge von Datensätzen formal zu bestimmen, sind **k-Anonymität** und darauf aufbauend **l-Vielfalt** und **t-Nachbarschaft**:

- **k-Anonymität**: Der Parameter k gibt an, wie leicht Datensätze aus einer Tabelle anhand der indirekten Identifier einer natürlichen Person zugeordnet werden können. Ist z.B. k = 5, so gehört jeder Datensatz zu einer Gruppe (sog. Äquivalenzklasse) von mindestens fünf Datensätzen, die in allen identifizierenden Merkmalen Attributen gleich sind (z.B. gleiches Alter, gleiche PLZ etc.). Je höher der k-Wert, umso schwerer ist es für einen Angreifer, einen ihm zugänglichen Datensatz einer bestimmten natürlichen Person zuzuordnen, da jeweils mindestens k mögliche Personen in Frage kommen.
  Bei der k-Anonymität werden zur Bestimmung von k nur die „Quasi-Identifikatoren“ (auch indirekte Identifier) betrachtet, nicht aber weitere „sensible Attribute“, die in der Tabelle vorhanden sind (z.B. Krankheitsbilder in einer Patientenakte). Bei den sensiblen Attributen wird vielmehr unterstellt, dass sie einem Angreifer nicht bekannt sind, sondern ihre Kenntnis gerade das Ziel des Angriffes darstellt. Die k-Anonymität einer Tabelle kann u.a. erhöht

werden, indem Informationen generalisiert (z.B. Altersgruppen statt Geburtsdatum) oder Attribute entfernt werden. Beispiel:

Rohdaten:

	Quasi-Identifikatoren			sensible Attribute	
	Alter	**Geschlecht**	**PLZ**	**Diagnose**	**Behandlung**
1	20	W	20089	Appendizitis	Operation
2	23	M	22527	Erkältung	Medikament XY
3	26	M	22550	Bruch	Gips
4	27	W	20055	Erkältung	Medikament XY
5	30	M	20089	Bruch	Schiene
6	34	M	20089	Bruch	Gips
7	39	M	20014	Appendizitis	Operation

Anonymisierung durch Generalisierung bzgl. Alter und PLZ:

	Quasi-Identifikatoren			sensible Attribute	
Gruppe	**Alter**	**Geschl.**	**PLZ**	**Diagnose**	**Behandlung**
1	20-29	M	225**	Erkältung	Medikament XY
	20-29	M	225**	Bruch	Gips
2	20-29	W	200**	Appendizitis	Operation
	20-29	W	200**	Erkältung	Medikament XY
3	30-40	M	200**	Bruch	Schiene
	30-40	M	200**	Bruch	Gips
	30-40	M	200**	Appendizitis	Operation

Die generalisierte Tabelle enthält drei Äquivalenzklassen (Gruppen) mit jeweils mindestens zwei Datensätzen. Sie ist damit 2-anonym.

Tabelle 1: k-Anonymität

- **l-Vielfalt (l-diversity)**: l-Vielfalt erweitert die k-Anonymität mit Blick auf „Homogenitätsattacken". Wenn innerhalb einer k-Äquivalenzgruppe ein sensibles Attribut für alle Datensätze gleich ist (zum Beispiel alle Patienten an derselben Krankheit leiden), dann weiß der Angreifer auch ohne Zuordnung zu einer bestimmten Person, dass dieses Attribut auf alle Personen der Äquivalenzgruppe zutrifft („Homogenitätsattacke"). Um dieses Risiko zu berücksichtigen, misst man die l-Vielfalt. Der Parameter l gibt an, dass ein sensibles Attribut innerhalb einer Äquivalenzgruppe mindestens l unterschiedliche Ausprägungen aufweist.
- **t-Nachbarschaft (t-closeness)**: Das Konzept der t-Nachbarschaft verbessert den Ansatz der l-Vielfalt und schließt weitere Möglichkeiten aus, um aus einer Tabelle sensible Informationen über Individuen abzuleiten. Insbesondere wird berücksichtigt, dass die Verteilung

innerhalb einer Äquivalenzgruppe von der Verteilung über die gesamte Tabelle mehr oder weniger stark abweichen kann und dies ggf. wichtige Informationen offenbart. Ist zum Beispiel ein Merkmal wie „HIV-positiv" über die gesamte Tabelle sehr selten (<1%), aber innerhalb einer bestimmten k-Äquivalenzgruppe häufig (z.B. 50%), dann kann aus der Tabelle abgeleitet werden, dass die Mitglieder dieser Gruppe eine vergleichsweise hohe Wahrscheinlichkeit für dieses Merkmal haben – allein dies ist bereits eine sehr sensible Information. Um solche Schlüsse zu erschweren, verlangt die t-Nachbarschaft, dass die Verteilung eines sensiblen Attributs innerhalb einer Äquivalenzklasse maximal um einen Parameter t von der Verteilung des Attributs über die gesamte Tabelle abweichen darf. Die Anwendung von t-Nachbarschaft garantiert somit besser als die Methode der l-Vielfalt, dass eine Tabelle nur einen minimalen Wissensgewinn über Individuen zulässt. Sie ist aber komplexer in der Anwendung, da zur Bestimmung von t zunächst ein Distanzmaß für die Merkmalsverteilungen definiert werden muss und dabei semantische Ähnlichkeiten zwischen den Merkmalsausprägungen zu berücksichtigen sind.

Die Parameter k, l und t liefern **formale Maße** für die Anonymität im Sinne der Möglichkeit, aus einer gegebenen Tabelle personenbezogene Informationen abzuleiten. Diese Maße berücksichtigen aber allein den **theoretischen Informationsgehalt**. Die Wahrscheinlichkeit der Identifizierung und die Frage, ob personenbezogene Daten vorliegen, ist anhand der weiteren zuvor skizzierten Kriterien zu bewerten und zu beantworten.[55] Zu berücksichtigen ist hierbei auch, dass die technischen Möglichkeiten (Rechengeschwindigkeit, Speicherplatz) sich stetig weiterentwickeln und damit in der Praxis die Wahrscheinlichkeit einer Identifizierung mit fortschreitender Zeit zunehmen kann.

5.1.2 Differential Privacy

Auch das Konzept der Differential Privacy zielt darauf ab, den Grad der Anonymität einer Datenbank messbar zu machen, allerdings in dem besonderen Fall, dass diese Datenbank vom Betreiber Dritten (z.B. Forschenden oder der Öffentlichkeit) über eine Schnittstelle für Abfragen bereitgestellt wird. Anders als die unter 5.1.1 genannten Anonymitätsmaße setzt die Differential

[55] Siehe oben 1.

Privacy daher nicht bei den in der Datenbank enthaltenen Daten an, sondern bezieht sich auf die Informationen, welche von den Dritten durch Abfragen (sog. „Queries") aus der Datenbank abgefragt werden können. Differential Privacy beschreibt einen Formalismus, mit dem sich bemessen lässt, mit welcher Wahrscheinlichkeit bzw. Konfidenz jemand durch Queries Informationen über Einzelpersonen aus der Datenbank extrahieren kann.
Ein System, das den Anforderungen von Differential Privacy entspricht, soll statistische Abfragen an eine Datenbank ermöglichen. Die Daten einzelner Menschen sollen dabei jedoch keinen Einfluss auf das Ergebnis der Abfragen haben. Es sollen also nur statistische, aber keine personenbezogenen Informationen abgefragt werden können (auch wenn in der Datenbank ggf. personenbezogene Daten enthalten sind).
Wenn die Datenbank die im konkreten Fall angestrebten Werte für Differential Privacy nicht von sich aus erfüllt, wird den Antworten auf Queries **Rauschen** hinzugefügt (siehe Randomisieren). Insofern kombiniert Differential Privacy ein Anonymitätsmaß mit einem operativen Ansatz, um das gewünschte Maß an Anonymität zu erreichen.
In der Praxis besteht die Herausforderung darin, das hinzugefügte Rauschen so einzurichten, dass ein vernünftiger Ausgleich zwischen Datenschutz einerseits und der Tauglichkeit des Systems für die verfolgten Zwecke andererseits besteht. Es sollten sich also trotz des Rauschens noch statistisch valide Aussagen ableiten lassen bzw. sollte – bei Einsatz von Machine Learning – es noch möglich sein, KI-Modelle erfolgreich auf den Daten zu trainieren.[56]
Anwendungsbeispiele von Differential Privacy sind:

- **Demografische Daten** (z.B. Zensus-Ergebnisse), die für statistische Abfragen über das Internet bereitgestellt werden
- Daten, die gemäß dem **Data-Governance-Act** durch öffentliche Stellen zur Weiterverwendung bereitgestellt werden[57]

[56] Vgl. Blanco-Justicia et. al., A Critical Review on the Use (and Misuse) of Differential Privacy in Machine Learning, ACM Computing Surveys 2022 (55), 1 ff.
[57] Vgl. Erwägungsgrund 7 Data-Governance-Act (VO (EU) 2022/868).

- Daten, die innerhalb der Verwaltung anderen Behörden zu **Analysezwecken** weitergegeben werden[58]
- **Patientenakten**, die zu Forschungszwecken für statische Datenbankabfragen oder für das Training von Machine-Learning-Modellen bereitgestellt werden (etwa um die Vorhersage, Erkennung oder Behandlung von Krankheiten zu verbessern oder die Wirksamkeit von Medikamenten zu bewerten).
- **Nutzungsdaten** von Internetdiensten, die mittels Differential Privacy anonymisiert öffentlich bereitgestellt werden (z.B. Statistiken von Facebook oder Wikimedia) oder Nutzungsdaten von IT-Endgeräten, die mit Differential Privacy lokal anonymisiert erfasst werden (z.B. Smartphones, Laptops von Unternehmen wie Apple, Microsoft).

5.2 Interaktive und nicht-interaktive Verfahren

Zu unterscheiden ist zwischen **interaktiven** und **nicht-interaktiven** Anonymisierungsverfahren. Bei nicht-interaktiven Verfahren werden die Daten bzw. die Datenbank selbst anonymisiert und im Anschluss daran veröffentlicht (oder bestimmten Empfängern übermittelt). Bei interaktiven Verfahren wird die Datenbank selbst nicht verändert, sie bleibt in den Händen des Herausgebers und wird den Empfängern lediglich über eine Schnittstelle zur Verfügung gestellt, über die sie Anfragen an die Datenbank schicken können; die Anonymisierung findet an der Schnittstelle statt, indem sichergestellt wird, dass die den Empfängern übermittelten Ausgabedaten anonym sind. Eine interaktive Anonymisierung findet etwa im Rahmen der „Differential Privacy" statt, bei der die Query-Ergebnisse verrauscht werden (s.o. 5.1.2).

Jedes interaktive Anonymisierungsverfahren kann auch nicht-interaktiv eingesetzt werden (zum Beispiel indem der Herausgeber der Datenbank selbst Abfragen über die anonymisierende Schnittstelle stellt und die Ergebnisse veröffentlicht). Umgekehrt können Methoden wie Löschen, Slicing, Maskieren etc., die vielfach nicht-interaktiv angewandt werden, prinzipiell auch in eine Abfrage-Schnittstelle eingebaut und damit als interaktive Verfahren eingesetzt werden. Das bekannteste interaktive Verfahren ist die Differential Privacy, die als Anonymisierungsmethode das Verrauschen einsetzt.

[58] Dies wird z.B. in der Schweiz aktiv erforscht und angewandt, siehe die Website des Bundesamts für Statistik: <https://www.bfs.admin.ch/bfs/de/home/dscc/blog/2024-05-dp.html>.

Interaktive Verfahren sind in der Regel aufwändiger als nicht-interaktive Verfahren, bieten aber eine höhere Kontrolle über die Nutzung der Daten und ermöglichen damit im Allgemeinen ein höheres Schutzniveau.

5.3 Anonymisierungstechniken im Überblick

Es gibt eine Vielzahl von Anonymisierungstechniken. Welche davon anwendbar sind, hängt von der Art der zu anonymisierenden Ausgangsdaten ab und davon, wie die anonymisierten Daten genutzt werden sollen.

Durch die Anonymisierung werden die Ausgangsdaten verändert (transformiert), so dass ein neuer, anonymisierter Datenbestand entsteht. Dabei kann unterschieden werden zwischen Methoden, bei denen die anonymisierten Daten informationsärmer, aber weiterhin korrekt sind („**truthful transformations**“) und Methoden, bei denen die anonymisierten Daten ggf. auch inhaltlich verfälscht werden („**deviations from the truth**“). Zur ersten Kategorie zählen das Löschen (5.3.1), Maskieren (5.3.2), Generalisieren (5.3.3), Aggregieren (5.3.4) und das Slicing (5.3.5). Zu den verfälschenden Anonymisierungstechniken zählen das Top- and Bottom-Coding (5.3.6), das Randomisieren (5.3.7) und das Generieren synthetischer Daten (5.3.8).

5.3.1 Nichtangabe/Löschen/Unterdrücken

Beim Löschen werden Teile der identifizierenden Merkmale (d.h. der Identifikatoren und der Quasi-Identifikatoren) aus den Datensätzen gelöscht. Die Datensätze enthalten somit nach dem Löschen weniger Informationen. Der nicht gelöschte Teil der Informationen bleibt unverändert. Beispiele für das Löschen sind:

- **Löschen** oder Kürzen von **Identifikatoren** aus Datensätzen (z.B. bei Schlüsselattributen wie Kundennummer, Personalnummer etc.)
- **Löschen** von **Quasi-Identifikatoren** aus Datensätzen, z.B.
    - Löschen einzelner Merkmale einzelner Personen (insbesondere Ausreißer, Extremwerte oder singuläre Werte können eine Identifikation erleichtern und kommen daher für eine Löschung in Frage; z.B. könnte die Größenangabe von besonders kleinen oder großen Menschen aus einer Tabelle gelöscht werden)
    - Löschen einzelner Merkmale für alle Personen (dies entspricht dem Löschen einer Spalte aus einer Tabelle und bietet sich an bei Merkmalen mit einer hohen Identifikationskraft wie z.B. einer genauen Geolokation)

- Löschen aller Merkmale eines einzelnen Betroffenen (kommt in Betracht, wenn einzelne Betroffene eine sehr seltene und identifizierende Merkmalskombination aufweisen)
- Löschen aller Merkmale einer Gruppe von Betroffenen (dies wird häufig genutzt, um eine bestimmte k-Anonymität sicherzustellen, indem Cluster gleichartiger Betroffener, die weniger als k Mitglieder haben, vollständig gelöscht werden – sog. „cell suppression")

- **Löschen** von Daten aus Zeitreihen (Verringerung der Auflösung durch „Resampling")
- **Löschen** von Anfangs- und Endpunkt einer Route, um bei Mobilitätsdaten die Identifizierung zu erschweren

Beim Löschen ist zu beachten, dass es einen erheblichen Unterschied machen kann, ob die Löschung eines Merkmals ausdrücklich sichtbar gemacht wird, also unterscheidbar ist von einem schlicht fehlenden Wert. Beispiel: Werden im Merkmal „Geschlecht" Werte, die weder männlich noch weiblich sind, durch ein „*" ersetzt statt durch einen Eintrag mit der Bedeutung „Angabe fehlt", so ermöglicht diese explizite Ausweisung Rückschlüsse auf ein sensibles Datum. Es handelt sich dann nicht um ein Löschen, sondern um eine Maskierung (dazu sogleich 5.3.2).

5.3.2 Maskieren/Ersetzung

Das Maskieren ähnelt dem Löschen, jedoch werden die gelöschten Daten durch Platzhalter ersetzt. Beispiele:

Praxishinweis: Das Maskieren von Identifikatoren und Quasi-Identifikatoren erfolgt z.B. bei der Veröffentlichung von Daten in Transparenzregistern oder aufgrund von IFG-Anfragen, indem bestimmte Angaben „geschwärzt" werden.

- **Maskieren** von Textdateien, z.B. durch
  - **Schwärzen** von Namen oder sonstigen Identifiern
  - **Ersetzen** von Namen oder sonstigen Identifiern durch Platzhalter („Aus-X-en", Überschreiben mit Sternchen); ggf. teilweises Ersetzen (z.B. der letzten drei Stellen einer IBAN oder der letzten beiden Oktette einer IPv4-Adresse)
- **Maskieren** von Bildern oder Videos, z.B. durch
  - **Verdecken** von Bildbereichen, die Gesichter, Kfz-Kennzeichen oder ähnliche Informationen enthalten (z.B. schwarze Balken in Fotos)
  - **Verfremden** von Bildbereichen (z.B. durch Hinzufügen von Unschärfe)

5.3.3 Generalisieren

Beim Generalisieren wird die Menge der möglichen Ausprägungen eines Merkmals auf eine kleinere Menge abgebildet. So werden mehrere Merkmalsausprägungen unter einer „unschärferen" Kategorie zusammengefasst.

Das Generalisieren kann auf Quasi-Identifikatoren und/oder auf sonstige sensible Merkmale angewandt werden. Durch das Generalisieren von Quasi-Identifikatoren fallen tendenziell mehr Betroffene in dieselben Cluster, sodass sich die k-Anonymität erhöht.

Beispiele für das Generalisieren sind:

- **Generalisieren von Altersangaben**
  - indem statt des genauen Geburtstags nur der Jahrgang angegeben wird (somit werden 365 mögliche Ausprägungen auf denselben Wert abgebildet) oder
  - indem mehrere Jahrgangskohorten in Töpfen/Intervallen zusammengefasst werden (z.B. 0-10 Jahre, 11-20 Jahre etc.) oder in Generationen
- **Generalisieren von Ortsangaben und Mobilitätsdaten**
  - indem statt genauer Geo-Koordinaten Zellen von z.B. 100m x 100m angegeben werden
  - indem mehrere Adressen zusammengefasst werden (z.B. „Mikrozelle" der Deutschen Post mit durchschnittlich 6,6 Haushalten)
- **Generalisieren von Zahlen durch Gruppieren** benachbarter Werte (z.B. Betrachten nur der ersten beiden Stellen einer PLZ) oder durch Runden bei Dezimalzahlen
- **Generalisieren sonstiger Werte durch Gruppieren** (z.B. Kategorie „europäisch" statt „deutsch"/„italienisch" etc.); die Gruppierung kann dabei ggf. auch über mehrere Merkmale hinweg erfolgen

5.3.4 Aggregieren

Eine besondere Form des Generalisierens ist das Aggregieren. Dabei werden die Datensätze mehrerer Personen zusammengefasst, um daraus eine Statistik zu bilden. Statistische Daten sind in der Regel nicht mehr personenbeziehbar, also vollständig anonym. Ausnahmefälle, in denen noch ein Personenbezug möglich sein kann bzw. zu prüfen ist, sind insbesondere:

- Aggregation jeweils nur sehr weniger Datensätze (z.B., wenn zu Controlling-Zwecken Kostenstellen gebildet werden, die aber jeweils nur die Ausgaben sehr weniger Beschäftigter zusammenfassen)

- Sehr informationsreiche Statistiken oder mehrere Statistiken, die kombiniert werden können, um Teile der Originaldatensätze zu rekonstruieren.

5.3.5 Slicing

Slicing zielt darauf ab, in hochdimensionalen Tabellen (also Tabellen mit vielen Spalten) strukturierte gespeicherte Daten auf bestimmte Weise aufzuteilen, um einerseits möglichst viel der vorhandenen Information zu erhalten, andererseits Rückschlüsse auf einzelne Individuen zu erschweren. Slicing kommt daher insbesondere bei hochdimensionalen Datensätzen in Betracht, die zum Beispiel für Forschungszwecke genutzt werden sollen.

Es gibt diverse Methoden des Slicings, die sich zum Teil erheblich unterscheiden und hier nicht im Detail dargestellt werden können. In der Regel wird beim Slicing zunächst – entsprechend dem Konzept der k-Anonymität – eine Gruppierung vorgenommen, sodass sich innerhalb der Gruppen die Individuen in einigen (oder allen) Quasi-Identifikatoren gleichen. Die Gruppen-ID wird dann der ursprünglichen Tabelle als neue Spalte hinzugefügt. Anschließend wird die Tabelle aufgeteilt, d.h. Spalten werden herausgenommen und als getrennte Tabelle gespeichert, wobei die Teil-Tabellen nur über die Gruppen-ID verbunden bleiben. Bei der Art und Weise der Aufteilung unterscheiden sich verschiedene Ansätze des Slicings. Beispiel:

Rohdaten:

	Quasi-Identifikatoren			
	Alter	Geschlecht	PLZ	Diagnose
1	20	W	20089	Appendizitis
2	23	M	22527	Erkältung
3	26	M	22550	Bruch
4	27	W	20055	Erkältung
5	30	M	20089	Bruch
6	34	M	20089	Bruch
7	39	M	20014	Appendizitis

Anonymisierung durch Slicing (Aufteilung):

Quasi-Identifikatoren			
Alter	Geschlecht	PLZ	Gruppe
23	M	22527	1
26	M	22550	1
20	W	20089	2
27	W	20055	2
30	M	20089	3
34	M	20089	3
39	M	20014	3

Gruppe	Diagnose	Anzahl
1	Erkältung	1
	Bruch	1
2	Appendizitis	1
	Erkältung	1
3	Appendizitis	1
	Bruch	2

Die Gruppen sind so gebildet wie bei der Generalisierung im Beispiel aus Ziffer 5.1.1. Beim Slicing werden die Quasi-Identifikatoren im Original mitgeliefert (mittlere Tabelle), aber von den sensiblen Merkmalen getrennt (untere Tabelle). Dies ermöglicht im Vergleich zur reinen Generalisierung weitergehende Auswertungsmöglichkeiten.

Tabelle 2: Slicing

5.3.6 Top- and Bottom-Coding

Oft sind Werte so verteilt, dass viele Individuen um die Mitte herum liegen und deutlich weniger an den Rändern. Dies ist zum Beispiel anschaulich bei Merkmalen wie der Körpergröße, die näherungsweise eine Gauß'sche Normalverteilung aufweisen – Fälle, die am oberen und unteren Rand liegen, sind selten und können damit Rückschlüsse auf einzelne Personen

ermöglichen. Um dies zu vermeiden, werden beim **Top- and Bottom-Coding** die zulässigen Wertebereiche oben und unten abgeschnitten. Im Beispiel der Körpergröße könnten (bei Erwachsenen) alle Werte <155cm als „160 cm" angegeben werden und alle Werte >195cm als „195 cm". Diese Transformation wird häufig für Statistiken genutzt. Sie führt allerdings (anders als die Methoden unter Ziffer 2.1-2.4) zu einer gewissen Verfälschung der Daten, die sich zum Beispiel auf die Berechnung von Durchschnittswerten auswirkt (so dass diese nicht mehr genau berechnet werden können).

5.3.7 Randomisieren

Daten können auch durch zufällige Veränderung der Werte anonymisiert werden (sog. Randomisierung). Die Daten werden dabei mehr oder weniger verfälscht und lassen insoweit keine sicheren Schlüsse mehr zu. Gängige Methoden sind:

- **Vertauschen/Permutation (Data Swapping)**: Beim Data Swapping werden einzelne Werte zwischen Datensätzen ausgetauscht. Typischerweise werden die Datensätze zuvor gruppiert (vgl. die Beispiele in Ziffer 5.1.1 und Ziffer 5.3.5) und Vertauschungen nur innerhalb der Gruppe vorgenommen. Das Vertauschen verfälscht zwar die betroffenen Datensätze, jedoch bleiben bestimmte wesentliche Merkmale der Datenbank erhalten (z.B. Verteilungen, Durchschnittswerte, Median).
- **Verrauschen/Noise Injection**: Beim Hinzufügen von Rauschen werden einzelne oder alle Datenpunkte zufällig verändert. Dies kann beispielsweise durch Addieren eines zufälligen Wertes („Störgröße") geschehen oder durch Ersetzen ausgewählter Werte mit Zufallswerten. Eine weitere Variante ist das Hinzufügen synthetischer Datensätze (siehe 5.3.8) zu einer Tabelle, sodass nicht mehr mit Sicherheit festgestellt werden kann, ob ein bestimmter Datensatz einer realen Person zugeordnet ist oder nicht („Dummy" bzw. „Fake").

Beispiel: Mobilitätsdaten, die zur Verkehrssteuerung erhoben werden (z.B. Standortdaten von Handys, Routen aus ÖPNV-Daten), bergen ein hohes Potential zur Identifizierung der Betroffenen. Durch Einstreuen von synthetisch generierten Routen von Fake-Verkehrsteilnehmern wird die Zuordnung erschwert und unsicher.

5.3.8 Synthetische Daten

Synthetische Daten sind Daten, die – anders als die korrespondierenden Echtdaten, die z.B. aus einer Messung stammen können – ohne Bezug zu realen Sachverhalten „künstlich" erzeugt wurden. Da synthetische Daten anstelle von Echtdaten eingesetzt werden sollen, müssen sie –

je nach Einsatzzweck – deren Eigenschaften bis zu einem gewissen Grad nachbilden (z.B. Struktur, Merkmalsverteilung). Folglich können synthetische Daten nicht vollkommen zufällig erzeugt werden, sondern bei der Generierung müssen die Eigenschaften der Echtdaten berücksichtigt werden. Hierbei ist darauf zu achten, dass keine personenbezogenen Rückschlüsse auf zugrundeliegende Echtdaten möglich sind. Ist dies gewährleistet, sind synthetische Daten vollständig anonym.

Wo der Einsatz synthetischer Daten möglich ist, ist dies durch den Grundsatz der Datensparsamkeit grundsätzlich geboten. Allerdings können nicht für alle Einsatzzwecke brauchbare synthetische Daten erzeugt werden. Z.B. können Softwaretests zwar im Allgemeinen mit synthetischen Daten durchgeführt werden, jedoch können damit zumeist nicht alle Fehler erkannt werden, da im Echtbetrieb Sonderfälle vorkommen können, die nicht vorhersehbar sind und bei der Erzeugung der synthetischen Daten nicht berücksichtigt wurden.

> Praxishinweis: Synthetische Daten können insb. zur ersten Testung von neuer Software eingesetzt werden.

5.3.9 Pseudonymisierung als „relative Anonymisierung"

Ausgehend vom relativen Verständnis des Personenbezugs ist eine (relative) Anonymisierung – d.h. eine Anonymisierung, die nicht gegenüber jeder denkbaren Stelle, aber gegenüber bestimmten Empfängern wirksam ist – möglich, und zwar u.a. durch **Pseudonymisierung**. Bei der Pseudonymisierung werden direkte Identifier durch Pseudonyme ersetzt; die pseudonymisierten Daten werden von den Zusatzdaten, welche eine Entschlüsselung der Pseudonyme ermöglichen (Zuordnungstabelle, Zuordnungsregel) organisatorisch getrennt. Einen Datenempfänger, der nur die pseudonymisierten Daten erhält, aber keinen Zugang zu den Zusatzdaten hat, sind die pseudonymisierten Daten dann anonym. Dies gilt jedoch nur unter der Annahme, dass dem Datenempfänger auch sonst keine Re-Identifizierung möglich ist (z.B. durch Verkettung mit weiteren Daten etc.).

Dasselbe gilt für **verschlüsselte Daten**. Soweit der jeweiligen datenverarbeitenden Stelle der Schlüssel fehlt, sind die Daten für sie anonym. Allerdings sind Daten in verschlüsselter Form regelmäßig nicht nutzbar. Daher gibt es nur wenige Dienstleistungen, die mit verschlüsselten Daten durchgeführt werden können, ohne dass der Dienstleister den Schlüssel hat (z.B. Backup-Service oder Kommunikationsdienste). Eine Ausnahme bildet die **homomorphe Verschlüsselung**, bei der bestimmte Auswertungen oder Berechnungen/Operationen auch mit den Daten in ihrer verschlüsselten Form möglich bleiben.

5.3.10 Treuhandlösungen

Bei der praktischen Durchführung einer Anonymisierung zur Bereitstellung von Daten an Dritte können Datentreuhänder und Datenvermittlungsdienste eine wichtige Rolle einnehmen. Diese können zum Beispiel Pseudonyme verwalten oder vergeben und dabei die Trennung der pseudonymisierten Daten von der Zuordnungsregel bei den Empfängern gewährleisten. Ebenso können Sie Daten anonymisieren oder Daten in sicheren Verarbeitungsumgebungen bereitstellen. Beispiele hierfür sind die Datenvermittlungsdienste nach dem Data Governance Act.[59]

5.4 Kombinieren von Anonymisierungstechniken

Anonymisierungsverfahren müssen auf den Zweck und die konkreten Umstände der Verarbeitung sowie die damit verbundenen Risiken zugeschnitten werden. In der Praxis bietet es sich oft an, mehrere Anonymisierungstechniken zu kombinieren, um einen optimalen Datenschutz zu erreichen und zugleich die Eignung der Daten für den verfolgten Zweck zu erhalten.

5.5 Anonymisierung bei mehreren Beteiligten

Im vorherigen Abschnitt sind technische Anonymisierungsverfahren behandelt worden, die auf die objektive Qualität von Daten abstellen. Aus rechtlicher Sicht kommt hinzu, dass der Personenbezug von Daten – und damit umgekehrt der Begriff der Anonymität – „relativ" verstanden wird, also aus der Perspektive der jeweiligen datenverarbeitenden Stelle. Somit gibt es aus rechtlicher Sicht Verfahren, die (nur) zu einer Anonymisierung gegenüber bestimmten Beteiligten führen (zum Beispiel gegenüber einem Auftragsverarbeiter).

Die relative Betrachtung des Personenbezugs ist auch bei der Beurteilung der Frage anzuwenden, ob trainierte KI-Modellen personenbezogene Daten enthalten.[60] Der Europäische Datenschutzausschuss geht davon aus, dass in KI-Modellen prinzipiell personenbezogene Daten gespeichert sein können, auch wenn diese nicht unmittelbar wahrnehmbar repräsentiert sind (sondern in Form von Gewichten bzw. Wahrscheinlichkeitswerten). Jedoch ist dies nicht

[59] Vgl. Art. 2 Nr. 10 und Art. 10 ff. Daten-Governance-Rechtsakt (VO (EU) 2018/1714).

[60] Der Begriff des KI-Modells wird in der KI-Verordnung nur näherungsweise in Erwägungsgrund 97 bestimmt. Danach sind KI-Modelle wesentliche Komponenten von KI-Systemen, stellen aber für sich genommen keine KI-Systeme dar. Damit KI-Modelle zu KI-Systemen werden, ist die Hinzufügung weiterer Komponenten, zum Beispiel einer Nutzerschnittstelle, erforderlich. Siehe dazu oben 2.5.

zwingend und auch für KI-Modelle ist eine datenschutzrechtlich anonyme Arbeitsweise denkbar.[61] Entscheidend ist, ob personenbezogene Daten tatsächlich aus dem Modell ableitbar sind (durch Identifizieren einzelner Betroffener, Verknüpfung mit anderen Daten oder Inferenz). Die Beurteilung des Personenbezugs für ein konkretes Modell basiert darauf, welche Mittel und Methoden der Verantwortliche oder eine andere Person wahrscheinlich einsetzen wird.[62] Mithin ist es möglich, dass die ein KI-System trainierende oder entwickelnde Stelle Möglichkeiten zur Ableitung personenbezogener Daten hat (z.B. aufgrund von Zusatzwissen und technischen Methoden), nicht aber die intendierte Gruppe der Nutzenden, so dass die Daten für die Nutzenden anonym sind. Ggf. kann auch im KI-System durch entsprechende Ausgabefilter sichergestellt werden, dass die Ausgabedaten anonymisiert sind (s. 5.2. zu interaktiven Anonymisierungsverfahren). Wird ein Modell nicht nur einer bestimmten Gruppe von Nutzenden bereitgestellt, sondern veröffentlicht, so ist bei der Betrachtung der ggf. zur Identifizierung eingesetzten Mittel allerdings konsequent auf alle denkbaren Stellen und deren Mittel und Zusatzwissen abzustellen.

> Praxishinweis: Wird in einer Behörde eine KI-Lösung eingekauft, ist sorgfältig zu prüfen, inwiefern den Nutzenden Daten aus dem KI-Modell zugänglich sind. Ist der Zugang wahrscheinlich, ist eine Rechtsgrundlage für die Datenverarbeitung notwendig.

5.6 Besondere Herausforderungen

Der erforderliche Aufwand für die Anonymisierung wächst mit der **Informationsdichte** der Daten. Das liegt daran, dass die Gefahr einer erfolgreichen Re-Identifizierung durch Singling-Out oder Verkettung der Daten mit Zusatzwissen mit steigender Informationsdichte zunimmt. Dieses Problem stellt sich insbesondere bei strukturierten Daten, bei denen für jeden Datenpunkt eine Vielzahl von Attributen vorliegt (**hochdimensionale Daten**). Auch unstrukturierte Daten weisen häufig eine hohe Informationsdichte auf und müssen daher besonders sorgfältig anonymisiert werden.

[61] *EDSA*, Opinion 28/2024, 17. Dezember 2024.
[62] *EDSA*, Opinion 28/2024, 17. Dezember 2024, Rn. 40 ff.

Abbildungsverzeichnis

Tabellenverzeichnis

Quellenverzeichnis

EuGH, Urteil vom 9. November 2023, Gesamtverband Autoteile-Handel eV ./. Scania CV AB, C-319/22, ECLI:EU:C:2023:837.

EuGH, Urteil vom 19. Oktober 2016, Patrick Breyer./. Bundesrepublik Deutschland, C-582/14, ECLI:EU:C:2016:779.

EuGH, Urteil vom 7. März 2024, IAB Europe, C-604/22, ECLI:EU:C:2024:214.

EDPB, Guidelines 01/2025 on Pseudonymisation, adopted on 16 January 2025.

*Fuchs, Thomas***,** Hamburger Thesen zum Personenbezug in LLMs Ein Debattenimpuls zur Anwendbarkeit der DS-GVO auf Large Language Models, KIR 2024, S. 79 ff.

EDSA, Opinion 28/2024, 17. Dezember 2024.

Hornung, Gerrit/Wagner, Bernd, Anonymisierung als datenschutzrelevante Verarbeitung?, ZD 2020, S. 223 ff.

Gola, Peter/Heckmann, Dirk, Datenschutz-Grundverordnung – Bundesdatenschutzgesetz, 3. Auflage 2022.

Schantz, Peter, Die Datenschutz-Grundverordnung – Beginn einer neuen Zeitrechnung im Datenschutzrecht, NJW 2016, 1841 ff.

Blanco-Justicia et. al., A Critical Review on the Use (and Misuse) of Differential Privacy in Machine Learning, ACM Computing Surveys, 2022.

Bundesamts für Statistik der Schweizerischen Eidgenossenschaft, Fragen und Antworten: Differential Privacy, abrufbar unter: https://www.bfs.admin.ch/bfs/de/home/dscc/blog/2024-05-dp.html, (zuletzt abgerufen am 28.4.2025).