Lernende Systeme
DIE PLATTFORM FÜR KÜNSTLICHE INTELLIGENZ

# Datenschatz für KI nutzen, Datenschutz mit KI wahren

## Whitepaper

Müller-Quade, J., Houdeau, D. et al.
AG IT-Sicherheit, Privacy, Recht und Ethik

## Kurzfassung

### Zusammenfassung

Künstliche Intelligenz bietet enormes Potenzial für innovative Geschäftsmodelle, die Stärkung unseres Zusammenlebens sowie für ökologische Nachhaltigkeit. Ihr Potenzial fußt auf einer stetig wachsenden Menge und Verfügbarkeit an Daten, darunter auch sensible Daten. Doch wie kann dieser Datenschatz in erfolgreiche KI-Anwendungen gebracht werden? Aktuell nutzen laut einer repräsentativen Studie nur 13 Prozent (Bitkom Research, 2021) der deutschen Unternehmen dieses Datenpotenzial. Grund sind die bestehenden datenschutzrechtlichen Bestimmungen, die in der rechtlichen Auslegung oft unsicher sind und damit die KI-Anwendung in der Breite erschweren. Zu diesem Spannungsfeld Datenschatz und Datenschutz liefert das Whitepaper eine lösungsorientierte Perspektive, die den Datenschutz und eine flexibilisierte Datennutzung zusammendenkt. Dies, indem es die gesetzliche Anerkennung technischer Privacy-Ansätze auf europäischer Werteordnung fordert, um die Rechtssicherheit für Unternehmen beim Einsatz und der Entwicklung von KI zu stärken. Voraussetzung: die Nutzung personenbezogener Daten ist alternativlos und liegt im Interesse des Gemeinwohls.

## Interaktionsfeld Datenschutz und gemeinwohlorientierte Datennutzung für KI-Systeme

Die Basis für gemeinwohlorientierte Errungenschaften stellt der Datenschatz dar, der beim Einsatz von KI-Technologien für das Training von **Machine-Learning Modellen** erforderlich ist: Gesundheitsdaten von Patientinnen und Patienten könnten beispielsweise dafür verwendet werden, die Detektion und Therapie von Erkrankungen zu erleichtern und zu optimieren; Bewegungsdaten von Personen und Fahrzeugen, um den Ressourcenverbrauch zu reduzieren. In der öffentlichen Verwaltung könnte durch die Automatisierung von Verwaltungsvorgängen mit personenbezogenen Daten der Arbeitsalltag zunehmend erleichtert werden. All diese Datenmengen sind bereits vorhanden und wachsen stetig, da unsere Lebenswirklichkeit zunehmend digitalisiert ist. Die Nutzung des Datenschatzes für KI-Systeme ist jedoch komplex, insbesondere bei personenbezogenen Daten. So stehen Unternehmen vor zahlreichen Herausforderungen im Umgang mit Datenschutzfragen zu personenbezogenen Daten: Einwilligung – ja oder nein? Ist sie notwendig? Wie lange ist diese aufzuheben? Wie sind personenbezogene Daten zu behandeln, die in die Kreislaufwirtschaft eines Unternehmens einfließen? Denn die Datenschutzgrundverordnung (DSGVO) stellt gerade an die Nutzung von personenbezogenen Daten hohe Anforderungen. In der Umsetzung der Datenschutzvorgaben im Rahmen der Datennutzung für KI zeigt sich zudem, dass die rechtliche Auslegung teils sehr unsicher ist. So treten durch alle Phasen des Daten-Lebenszyklus hindurch von der Datenerhebung, -prozessierung, -analyse, über -veröffentlichung, -speicherung bis hin zur -sekundärnutzung rechtliche Unsicherheiten in der Anwendung auf. Dies lässt viele Unternehmen vor dem Einsatz und der Entwicklung von KI-Systemen zurückschrecken, sodass das Potenzial des Datenschatzes nicht voll ausgeschöpft wird. Um diese aktuell bestehenden Interpretationsspielräume bei der Verarbeitung personenbezogener Daten schließen zu können, sind anwendungsspezifische Handlungsräume seitens der Gesetzgeber zu schaffen und technische Verfahren zur Wahrung des Datenschutzes juristisch zuzulassen. Dies erfordert einerseits technische Innovationen zur Minimierung der Rechtsunsicherheiten, andererseits gezielte, anwendungsspezifische Gesetzgebung.

**Abbildung: Datenschutzauflagen an KI-Systeme und Interpretation in der Anwendung über den Daten-Lebenszyklus**

| PHASE IM DATEN-LEBENSZYKLUS | DATENSCHUTZ – RECHTLICHE VORGABEN | UNSICHERHEIT IN DER ANWENDUNG |
|---|---|---|
| DATEN-ERHEBUNG | · Einwilligung der betroffenen Personen (Art. 4 Nr. 11 DSGVO)<br>· Informiertheit der betroffenen Personen (Art. 6 Abs. 1 lit. a) DSGVO)<br>· Zweckbindung der Datenerhebung (Art. 5 Abs. 1 lit. b) DSGVO) | · Wie granular/detailliert muss die Einwilligung sein?<br>· Welcher Grad an Informiertheit ist erforderlich?<br>· Wie soll Zweckbindung interpretiert werden? |
| DATEN-AUFBEREITUNG | · Grundsatz der Datenminimierung (Art. 5 Abs.1 lit. c) DSGVO) | · Datenminimierung erschwert Sekundärnutzung der Datensätze |
| DATEN-ANALYSE | Bei automatisierten Entscheidungen gilt<br>· Mitteilungspflicht nach Art. 23 Abs. 2 lit. f) und Art. 14 Abs. 2 lit. g) DSGVO, sobald Personen von automatisierten Entscheidungen nach Artikel 22 DSGVO betroffen sind<br>· Auskunftsrecht nach Art. 15 Abs. 1 lit. h) DSGVO<br>· Widerspruchsrecht nach Art. 21 Abs. 1 DSGVO gegen Datenverwendung sicherstellen | · Ab wann fällt ein KI-System unter die Rechtsdefinition „automatisierte Entscheidungen"? |
| DATEN-VERÖFFENTLICHUNG | · Zweckbindungsgrundsatz (Art. 5 Abs. 1 lit. b) DSGVO)<br>· Rechtmäßigkeit der Verarbeitung (Art. 6 Abs. 1 DSGVO)<br>· Jederzeit widerrufbare Einwilligungserklärung der Betroffenen (Art. 6 Abs. 1 lit. a) DSGVO) | |
| DATEN-SPEICHERUNG | · Aufbewahrungs-/Löschfristen (Art. 17 DSGVO)<br>· Identifizierung betroffener Personen nur so lange erlaubt, wie es ursprünglicher Datenverarbeitungszweck erfordert | · Interpretationsspielraum in der Rechtsauslegung<br>· Uneinheitlichkeit bei Aufbewahrungs-/Löschfristen<br>· Schwierige Anonymisierbarkeit von Randgruppen<br>· Semantische Segmentation schafft Personenbeziehbarkeit |
| DATEN-SEKUNDÄRNUTZUNG | · Datensekundärnutzung ohne Einwilligung der Betroffenen erlaubt, wenn sie als „logischer nächster Schritt" gilt oder im Gemeinwohlinteresse liegt (Art. 6 Abs. 1 S. 1 lit. f) DSGVO) | · „logischer nächster Schritt" bietet Interpretationsspielraum<br>· Gemeinwohlinteresse ist nicht eindeutig definiert |

## Technische Ansätze zur datenschutzwahrenden gemeinwohlorientierten Datennutzung

Das Spannungsfeld Datenschutz und Datennutzung als symbiotische Verbindung zusammenzudenken schafft die Voraussetzung für eine flexibilisierte, datenschutzwahrende Datennutzung für und mit KI im Gemeinwohlinteresse. Hierbei können verschiedene technische Ansätze entlang des Daten-Lebenszyklus zum Einsatz kommen.

Über **Anonymisierungs- und Pseudonymisierungsmaßnahmen** kann bereits bei der Datenerhebung die Personenbezogenheit von Daten eliminiert werden, sodass die DSGVO nicht mehr greifen würde. Das Verfahren der Anonymisierung verspricht eine vollständige Eliminierung, während Pseudonymisierung weiterhin eine gewisse Personenbezogenheit beibehält. Beide technische Verfahren reduzieren jedoch erheblich die Datenqualität. Die

Datensynthese hingegen erzeugt synthetische Datensätze, die keine Personenbezogenheit aufweisen und nach geltender Rechtsauffassung nicht unter die DSGVO fallen und somit öffentlich geteilt werden können. Allerdings ist ihre Praktikabilität begrenzt: hoher Erzeugungsaufwand; nicht für Branchen geeignet, die für das Training ihrer ML-Modelle auf statistische Merkmale angewiesen sind, die eine persönliche Identifizierbarkeit ermöglichen. Das Verfahren Differential Privacy dagegen „verrauscht" Daten, sodass individuelle Datenpunkte nicht mehr identifizierbar sind. Dies ist insbesondere relevant für die Sicherung von Trainingsdaten. Allerdings ist bei Differential Privacy die praktische Umsetzung in Echtzeit und bei kleinen Datensätzen problematisch.

Auch **kryptografische Maßnahmen** bieten verschiedene Möglichkeiten zur Wahrung des Datenschutzes. Sie setzen im Daten-Lebenszyklus bei der Datenprozessierung an, um die Personenbezogenheit zu eliminieren. Aber aufgrund der hohen Rechenlast für die Bereitstellung für das Modelltraining sind sie schwächer hinsichtlich Skalierung. Als Verschlüsselungsmaßnahmen gewährleisten Homomorphic Encryption und Secure Multiparty Computation zwar die Input und Output Privacy, können jedoch aufgrund von Rechenlast und Kommunikationskosten in vielen Anwendungsbereichen unpraktikabel sein. Confidential Computing als weiteres kryptografisches Verfahren schließt eine Verschlüsselungslücke im Cloud Computing und ermöglicht die sichere Verarbeitung von Daten in einer vertrauenswürdigen Umgebung. Dies stärkt das Vertrauen in den Datenschutz, erfordert aber ebenfalls Kommunikationskosten.

Bei der Datenanalyse für das konkrete Training von KI-Modellen rücken vor allem **dezentrale ML-Methoden (Machine Learning)** in den Fokus, um Datenschutz und gemeinwohlorientierte Datennutzung zu kombinieren. Beim verteilten ML erfolgt das Training von ML-Modellen nicht auf einem zentralen Server, sondern auf den Endgeräten der datengebenden Personen. Damit verspricht ihr Einsatz hohe Datenqualität sowie hohe individuelle Datensouveränität und ist vor allem für KI-Entwickelnde interessant. Hierzu zählen Verfahren des verteilten maschinellen Lernens und hybride Ansätze verteilten maschinellen Lernens. Letztere versprechen, neue Angriffsvektoren schließen zu können, die Performance zu erhalten und gleichzeitig Datenschutz und Sicherheit zu gewährleisten.

Neben den zentralen wie dezentralen ML-Methoden könnten auch Verfahren wie **Erklärbare KI (XAI) oder Safe AI** als begleitende Optionen herangezogen werden, die die Entscheidungen und Ergebnisse von KI-Modellen für den Anwender besser interpretierbar bzw. nachvollziehbar machen sollen („painting the black box white"), und so die Akzeptanz für die Nutzung personenbezogener Daten für KI-Systeme steigern.

Einen innovativen Ansatz für eine datenschutzkonforme Datennutzungsflexibilisierung bieten auch **Datenzugriff-Managementsysteme** wie Datentreuhänder oder Personal Information Management System (PIMS), die zudem mehrere Phasen des Daten-Lebenszyklus abdecken. Ihr Ansatz für eine datenschutzkonforme Datennutzung basiert auf einem integrierten Datenmanagement innerhalb der Datenökonomie, wodurch Betroffenen eine aktivere und gleichberechtigtere Rolle an ihrer Datenverwendung für die KI-Entwicklung bzw. -Anwendung zukommt.

Die verschiedenen technischen wie nicht KI-Modell-bezogenen technischen Maßnahmen bieten über die einzelnen Phasen im Daten-Lebenszyklus die Möglichkeit für eine datenschutzkonforme Datennutzung für KI-Systeme im Gemeinwohlinteresse. Vor dem Hintergrund ihrer unterschiedlichen Leistungsfähigkeiten hinsichtlich Genauigkeit, Kommunikationsaufwand, Rechenlast, Datenmanagement sowie Invasivität bezogen auf die

Datenintegrität ist eine sorgfältige Abwägung und Auswahl der geeigneten Maßnahme im jeweiligem Anwendungskontext vorzunehmen. Entscheidend ist aber vor allem, dass diese Maßnahmen rechtlich anerkannt werden, um bestehende Unsicherheiten in der Auslegung der DSGVO aufzulösen.

## Gestaltungsoptionen und Ausblick

Für einen modernen und funktionsfähigen Rechtsrahmen hinsichtlich einer flexibilisierten datenschutzwahrenden Datennutzung im Gemeinwohlinteresse gilt sowohl im Allgemeinen als auch im Speziellen, dass er diejenigen Instrumente zulässt und fördert, die zur Eindämmung der Unsicherheit in der Rechtsauslegung beitragen und gleichzeitig kollidierende Rechte und Rechtsgüter umfassend schützt. Dies auf Grundlage eines **technikneutralen Datenschutzrechtsrahmens**, der anwendungsorientierte und klar interpretierbare Vorgaben spezifisch für die Datennutzung für KI-Systeme definiert, um so Rechts- und Handlungssicherheit für die KI-Entwicklung zu schaffen. Dieser technikneutrale Datenschutzrechtsrahmen basierend auf einem holistischen Ansatz sowie das Ermöglichen von Anpassungen bei technischen Maßnahmen im Sinne von „Privacy by Design" sollten zudem mit weiteren Anpassungen zur Datennutzung für KI-Technologien bei nachgewiesenem Gemeinwohlinteresse angereichert werden. Eine **einheitliche, breit angelegte Definition von Gemeinwohlinteresse** wäre grundlegend, um entlang konkreter Anwendungskontexte mit einem klar festgelegten Kriterienrahmen Gemeinwohlinteresse rechtssicher definieren zu können und im Bedarfsfall Datennutzungsflexibilisierungen zu ermöglichen. Essentiell ist hierfür eine grundsätzliche **rechtliche Anerkennung für eine flexibilisierte Datennutzung** im Gemeinwohlinteresse im bestehenden Datenschutzrechtsrahmen, die zugleich die **Rechtsfolgen der Flexibilisierung definiert**, um Handlungssicherheit für Datenverarbeitende und Orientierung für Betroffene zu gewährleisten. Um die Datennutzungsflexibilisierung zu unterstützen, sollten **technische Maßnahmen zur Anonymisierung personenbezogener Daten daher standardisiert und zertifiziert werden.** Vorzugsweise sollte die Verfügbarkeit von nicht-personenbezogene Daten gestärkt werden. Andernfalls, sofern die Verwendung personenbezogener Daten alternativlos ist, bedarf es einer **Stärkung und rechtlichen Anerkennung von Privacy by Design, der Datensouveränität wie der Datenkompetenzen datengebender Betroffener selbst.**

Die Ergebnisse des analytischen Teils wurden zudem in drei Anwendungsszenarien aus dem Bereich Mobilität (Anwendungsszenario LEASYNG), Bildung (Anwendungsszenario learn.digital) und Gesundheit (Anwendungsszenario vAItality) aufbereitet, um die zentrale Botschaft in kompakter Form exemplarisch aufzuzeigen: Datenschutz und eine flexibilisierte Datennutzung müssen zusammengedacht werden, um rechtssicheren Handlungsspielraum für Unternehmen zu gewährleisten!

**Impressum**
Herausgeber: Lernende Systeme – Die Plattform für Künstliche Intelligenz | Geschäftsstelle | c/o acatech | Karolinenplatz 4 | D-80333 München | kontakt@plattform-lernende-systeme.de | www.plattform-lernende-systeme.de | Folgen Sie uns auf X: @Lernende Systeme | LinkedIn: de.linkedin.com/company/plattform-lernende-systeme | Mastodon: social.bund.de/@LernendeSysteme | Stand: Oktober 2023 | Bildnachweis: sdecoret/Adobe/Titel

Diese Kurzfassung entstand auf Grundlage des Whitepapers *Datenschatz für KI nutzen, Datenschutz mit KI wahren. Technische und rechtliche Ansätze für eine datenschutzkonforme, gemeinwohlorientierte Datennutzung*, München, 2023. Es wurde erstellt von Mitgliedern der Arbeitsgruppe AG IT-Sicherheit, Privacy, Recht und Ethik. *https://doi.org/10.48669/pls_2023-5*

GEFÖRDERT VOM

Bundesministerium für Bildung und Forschung

acatech
DEUTSCHE AKADEMIE DER TECHNIKWISSENSCHAFTEN