

# Französische Datenschutzaufsicht definiert Möglichkeiten für die Forschungsdatennutzung - auch für die deutsche Forschung relevant

Am 19. Juni 2025 veröffentlichte die CNIL (Commission Nationale de l'Informatique et des Libertés) eine [Orientierungshilfe zur datenschutzkonformen Entwicklung von KI-Systemen](#) nach **Art. 6 Abs. 1 lit. f DSGVO** und nennt hierbei zahlreiche Anwendungsbeispiele in der Forschung.

Die CNIL stellt drei Bedingungen für ein berechtigtes Interesse des Verantwortlichen oder einer Dritten an der Datenverarbeitung auf:

## 1. Legitimität des Interesses

Das Interesse muss rechtmäßig, klar und präzise bestimmt sowie real und gegenwärtig sein

CNIL-Beispiele für ein legitimes Interesse: Durchführung wissenschaftlicher Forschungsarbeiten (insb. ohne Auftrag im öffentlichen Interesse); kommerzielle Interessen (sofern sie nicht gegen das Gesetz verstoßen und die Verarbeitung notwendig und verhältnismäßig ist - *EuGH, Urteil v. 4. Oktober 2024, Az.: C-621/22 - Tennisbond*)

## 2. Erforderlichkeit/Notwendigkeit der Verarbeitung für das verfolgte berechtigte Interesse

Die Verarbeitung muss die Erreichung des Interesses ermöglichen und es darf kein milderes Mittel existieren. Dabei ist der Grundsatz der Datenminimierung zu beachten.

## 3. Einzelfallabwägung: Keine unverhältnismäßige Beeinträchtigung der Grundrechte und Grundfreiheiten der betroffenen Personen

Je größer die antizipierten Vorteile der KI-Systeme (z.B. im Bereich der Gesundheitsfürsorge, Zugang zu grundlegenden Dienstleistungen), desto wahrscheinlicher ist das Überwiegen des legitimen Interesses an der KI-Datenverarbeitung ggü. den Rechten der betroffenen Person.

Dies steht in Abwägung mit der potenziellen Beeinträchtigung von betroffenen Personen durch die KI-Entwicklung bzw. -Nutzung, insbesondere mit Blick auf besonderes schutzwürdige, sensible Daten nach Art. 9 Abs. 1 DSGVO (z.B. biometrische Daten, genetische Daten und Gesundheitsdaten)

## CNIL-Maßnahmenkatalog an Datenschutzvorkehrungen:

- **Einschränkung der Erhebung oder Speicherung von personenbezogenen Daten**  
Anonymisierung bzw. Pseudonymisierung; bevorzugte Verwendung synthetischer Daten bei gleicher Leistungsfähigkeit des KI-Modells; Entfernung seltener oder abweichender Daten; Deduplizierung von Lerndaten; Verringerung des Verhältnisses der Anzahl der Modellparameter zum Volumen der Trainingsdaten: Regulierung der Zielfunktion des Trainings; Vertraulichkeitsgewährleistende Lernalgorithmen; Begrenzung des Überlernens; Sicherheitsmaßnahmen zur Verhinderung von Angriffen; etc.
- **Betroffenen Kontrolle über eigene Daten ermöglichen** (über die Anforderungen der DSGVO hinausgehend): erhöhte Transparenzpflichten; einfache und vorherige Ausübung des Widerspruchsrechts (bspw. Opt-Out oder Blacklist-Mechanismen); Recht auf Löschung; Ausübung der Betroffenenrechte aus der DSGVO erleichtern (bspw. durch Anwendungsprogrammierungsschnittstellen); Informationen bereitstellen; etc.
- **Begrenzung der Risiken in der Nutzungsphase:** technische Maßnahmen ergreifen; Reidentifizierungsmöglichkeit einschränken durch Lizenzen; Qualität der Lerndatenbank gewährleisten um ethische Diskriminierung zu vermeiden
- Weitere mögliche Maßnahmen: **Ernennung eines Ethikreferenten** (je nach Größe und Ressourcen des KI-Betreibers)