

BERECHTIGTES DATENSCHUTZRECHTLICHES INTERESSE AN DER VERARBEITUNG VON KI-TRAININGSDATEN

Markus Schröder, LL.M.

Hochschule Heilbronn

Herbstakademie 2025

Agenda

1. Einleitung – unterschiedliches Rechtsverständnis USA/EU
2. Rechtsverständnis USA
3. Rechtsverständnis EU
4. EDPB Opinion 28/2024
5. Vom EDPB offen gelassene Fragen
6. Rechtsprechung: OLG Köln, OLG SH
7. Fazit
8. Ausblick

Einleitung – unterschiedliches Rechtsverständnis



Einleitung – Rechtsverständnis USA

- ▶ **„reasonable expectation of privacy“**
 - ▶ 1: Subjektive Erwartungen des Betroffenen
 - ▶ 2: Objektive Erwartungen der Gesellschaft
 - ▶ s. aber auch Rspr. des **EGMR**
 - ▶ und **ErwG 47**

- ▶ **„publicly available information“ ->**

Einleitung – Rechtsverständnis USA

- ▶ **California Privacy Rights Act:** „information that a business has a reasonable basis to believe is **lawfully made available to the general public** by the consumer **or from widely distributed media**, or by the consumer; or information made available by a person to whom the consumer has disclosed the information if the consumer has not restricted the information to a specific audience.”

Einleitung – Rechtsverständnis USA

- ▶ **Texas Data Privacy and Security Act:** „information that a business has a reasonable basis to believe is **lawfully made available to the general public through widely distributed media**, by a consumer, or by a person to whom a consumer has disclosed the information, unless the consumer has restricted the information to a specific audience.”

Einleitung – Rechtsverständnis USA

- ▶ **Aber auch:**
 - ▶ P.M. et al v. OpenAI LP et al
 - ▶ A.T. & J.H. v. OpenAI LP et al

Einleitung – Rechtsverständnis EU

- ▶ Art. 9 Abs. 2 lit. e) DS-GVO und
- ▶ Art. 14 DS-GVO
- ▶ Grundsätzlich aber: **Art. 6 DS-GVO**

Art. 6 DS-GVO

- ▶ Art. 6 Abs. 1 a) (Einwilligung)
- ▶ Art. 6 Abs. 1 b) (Vertrag)
- ▶ **Art. 6 Abs. 1 f) (Berechtigte Interessen)**
- ▶ **Art. 6 Abs. 4 (Weiterverarbeitung)**

EDPB Opinion 28/2024

- ▶ „**The more precisely an interest is defined** in light of the intended purpose of the processing, the better it will enable to clearly apprehend the reality of the **benefits and risks** to be taken into account in the balancing test.” (Rn. 78)
- ▶ Depending on the case at stake, **there may also be risks to other fundamental rights**. For example, large-scale and indiscriminate data collection by AI models in the development phase may create a sense of surveillance for data subjects, **especially considering the difficulties to prevent public data from being scraped.**” (Rn. 80) ->

EDPB Opinion 28/2024

- ▶ „The impact of the processing on the data subjects may be influenced by (i) the **nature** of the data processed by the models; (ii) the **context** of the processing; and (iii) the **further consequences** that the processing may have.” (Rn. 83)

EDPB Opinion 28/2024

- ▶ **„Reasonable expectations play a key role in the balancing test**, not least due to the complexity of the technology used in AI models and the fact that it may be difficult for data subjects to understand the variety of potential uses of an AI model and the data processing involved. To this end, **the information provided to data subjects may be considered to assess whether data subjects can reasonably expect their personal data to be processed.**” (Rn. 92)

->

EDPB Opinion 28/2024

- ▶ „However, while the omission of information can contribute to the data subjects not expecting a certain processing, the mere fulfilment of the transparency requirements set out in the GDPR is not sufficient in itself to consider that the data subjects can reasonably expect a certain processing. Further, simply because information relating to the development phase of an **AI model is included in the controller’s privacy policy, it does not necessarily mean that the data subjects can reasonably expect it to happen**; rather, this should be analysed by SAs on the specific circumstances of the case and considering all of the relevant factors.” (Rn. 92)

->

EDPB Opinion 28/2024

- ▶ Bsp. **xAI**: „We may use your personal information for a variety of purposes. (...) To develop and improve our Service and to conduct research: For example to develop new product features, **to train our models**, to identify usage trends, to operate and expand our business activities, to identify new customers, and for data analysis.”
- ▶ Bsp. **Doctolib**: “**Trainieren von KI-Modellen**”; “**Berechtigtes Interesse** von Doctolib SAS als Verantwortlicher der Verarbeitung, wenn keine Gesundheitsdaten betroffen sind.“; „Ausdrückliche **Einwilligung** des Nutzers bei Verwendung personenbezogener Gesundheitsdaten.“ ->

EDPB Opinion 28/2024

- ▶ Datenschutzinformationen **können** vernünftige Erwartungen der betroffenen Personen begründen bzw. erschüttern
 - ▶ Vgl. **FTC**: Privacy Notices begründen bzw. erschüttern
reasonable expectation of privacy
- ▶ Allerdings erst ab dem Zeitpunkt gegebener Informationen
- ▶ KI-Training ohne gegebene Datenschutzinformationen dürfte jedenfalls auf Grundlage eines berechtigten Interesses unzulässig sein
- ▶ Zulässige Weiterverarbeitung? Art. 13 Abs. 3 DS-GVO?

EDPB Opinion 28/2024

- ▶ „**This Opinion does not analyse** the below provisions, which may still play an important role when assessing the data protection requirements applicable to AI models:”
 - ▶ “Processing of **special categories of data**”
 - ▶ “**Compatibility** of purposes”
 - ▶ “**Data protection impact assessments**”
 - ▶ “Principle of **data protection by design**”

->

Processing of special categories of data: EuGH

- ▶ **Art. 9 Abs. 2 e) DS-GVO:** „die Verarbeitung bezieht sich auf personenbezogene Daten, die die betroffene Person **offensichtlich öffentlich gemacht hat**“
- ▶ **Aber: EuGH - Rs. C-446/21:** „ist dahin auszulegen, dass der Umstand, dass sich eine Person bei einer öffentlich zugänglichen Podiumsdiskussion zu ihrer sexuellen Orientierung geäußert hat, dem Betreiber einer Onlineplattform für ein soziales Netzwerk **nicht gestattet**, andere Daten über die sexuelle Orientierung dieser Person zu verarbeiten, **die er gegebenenfalls außerhalb dieser Plattform** von Anwendungen und Websites dritter Partner im Hinblick darauf **erhalten hat**, sie zu aggregieren und zu analysieren, um dieser Person personalisierte Werbung anzubieten.“

Compatibility of purposes: Contextual Integrity

- ▶ „**Zusammenhang**, in dem die personenbezogenen Daten erhoben wurden, insbesondere hinsichtlich des Verhältnisses zwischen den betroffenen Personen und dem Verantwortlichen“ (**Art. 6 Abs. 4 b) DS-GVO = Kontext**)
- ▶ **Reasonable expectation vs. Contextual integrity (Helen Nissenbaum):**
- ▶ Kontext bleibt regelmäßig gewahrt, wenn der Verantwortliche, mit dem die Daten ursprünglich geteilt wurden, diese weiterverarbeitet

Compatibility of purposes: LfDI BW

- ▶ Berechtigtes Interesse (-), dann: **Art. 6 Abs. 4 DS-GVO**
- ▶ **a:** jede **Verbindung zwischen den Zwecken**, für die die personenbezogenen Daten erhoben wurden, und den Zwecken der beabsichtigten Weiterverarbeitung,
- ▶ **b:** den **Zusammenhang**, in dem die personenbezogenen Daten erhoben wurden, insbesondere hinsichtlich des Verhältnisses zwischen den betroffenen Personen und dem Verantwortlichen,

->

Compatibility of purposes: LfDI BW

- ▶ **c:** die Art der personenbezogenen Daten, insbesondere ob **besondere Kategorien personenbezogener Daten** gemäß Artikel 9 verarbeitet werden oder ob personenbezogene Daten über strafrechtliche Verurteilungen und Straftaten gemäß Artikel 10 verarbeitet werden,
- ▶ **d:** die möglichen **Folgen der beabsichtigten Weiterverarbeitung** für die betroffenen Personen,
- ▶ **e:** das **Vorhandensein geeigneter Garantien**, wozu Verschlüsselung oder **Pseudonymisierung** gehören kann.

Data protection impact assessments

- ▶ “in cases of likely high risk to the rights and freedoms of data subjects, Article 35 GDPR” (Rn. 56)
- ▶ Aber auch: “Accountability principle (Article 5(2) GDPR)” (Rn. 62)
- ▶ Art. 35 Abs. 1 Satz 1 DS-GVO: “Hat eine Form der Verarbeitung, insbesondere bei **Verwendung neuer Technologien**, aufgrund der Art, des **Umfangs**, der Umstände und der Zwecke der Verarbeitung voraussichtlich ein hohes Risiko für die Rechte und Freiheiten natürlicher Personen zur Folge, so führt der Verantwortliche vorab eine Abschätzung der Folgen der vorgesehenen Verarbeitungsvorgänge für den Schutz personenbezogener Daten durch.
- ▶ = (freiwillige) DSFA?

Principle of data protection by design: HBfDI

- ▶ **„Vorhandensein geeigneter Garantien (Pseudonymisierung)“:**
- ▶ **Hamburgischer Beauftragter für Datenschutz und Informationsfreiheit** (Diskussionspapier: Large Language Models und personenbezogene Daten):
- ▶ **Die bloße Speicherung eines LLMs stellt keine Verarbeitung im Sinne des Art. 4 Nr. 2 DSGVO dar. Denn in LLMs werden keine personenbezogenen Daten gespeichert.** Soweit in einem LLM-gestützten KI-System personenbezogene Daten verarbeitet werden, müssen die Verarbeitungsvorgänge den Anforderungen der DSGVO entsprechen. Dies gilt insbesondere für den Output eines solchen KI-Systems.“

OLG Köln, Urt. v. 23.05.2025 – 15 UKI 2/25

- ▶ Risikobasierte Prüfung der Interessenabwägung
- ▶ Aber: **EuGH - C-394/23**:
- ▶ Rn. 60: „Das berechtigte Interesse an kommerzieller Direktwerbung kann im Fall der Gefahr einer Beeinträchtigung der Grundrechte und Grundfreiheiten der betroffenen Person jedenfalls nicht überwiegen. Wie sich nämlich aus dem 75. Erwägungsgrund der DSGVO ergibt, können Risiken für die Rechte und Freiheiten natürlicher Personen – mit unterschiedlicher Eintrittswahrscheinlichkeit und Schwere – aus einer Verarbeitung personenbezogener Daten hervorgehen, die zu einem physischen, materiellen oder immateriellen Schaden führen könnte, insbesondere wenn eine solche Verarbeitung zu einer Diskriminierung führen kann.“ ->

OLG Köln, Urt. v. 23.05.2025 – 15 UKI 2/25

- ▶ Rn. 61: „In diesem Zusammenhang wird das vorlegende Gericht insbesondere zu prüfen haben, ob die von Mousse geltend gemachte Gefahr einer Diskriminierung aufgrund der Geschlechtsidentität besteht, und zwar vor allem im Licht der Richtlinie 2004/113, mit der der Grundsatz der Gleichbehandlung von Männern und Frauen beim Zugang zu und bei der Versorgung mit Gütern und Dienstleistungen verwirklicht wird.“
- ▶ (P) EuGH kann so verstanden werden, dass ein berechtigtes Interesse nicht überwiegen kann, soweit für die Rechte und Freiheiten der betroffenen Person (irgend-) ein Risiko besteht, ungeachtet der Eintrittswahrscheinlichkeit und Schwere des Risikos

OLG Köln, Urt. v. 23.05.2025 – 15 UKI 2/25

- ▶ Verweis auf KI-VO
- ▶ ErwG 105 Satz 2: „Für die Entwicklung und das Training solcher Modelle ist der Zugang zu riesigen Mengen an Text, Bildern, Videos und anderen Daten erforderlich.“
- ▶ Art. 2 Abs. 7 KI-VO?
- ▶ Aber Berücksichtigung bei Interessenabwägung?

OLG Schleswig-Holstein, Urt. v. 12.08.2025 – 6 UKI 3/25

- ▶ Keine Eilbedürftigkeit
- ▶ Jedoch tendenziell strenger Maßstab:
- ▶ Rn. 52: „(...) Maßnahmen wie De-Identifizierung und Tokenisierung (...) jedoch bereits nach eigenem Vorbringen keineswegs wirksam (...)“
- ▶ Aber auch Rn. 69: „Die Tatsache, dass die Verfügungsklägerin mit dem Ergebnis der Prüfung durch das OLG Köln nicht einverstanden ist, lässt weder eine Dringlichkeit neu entstehen noch lässt es die Notwendigkeit erkennen, zur Gewährung effektiven Rechtsschutzes zugunsten der Verbraucher trotz des erheblichen Zeitablaufs und der bereits erfolgten Datennutzung ein weiteres einstweiliges Rechtsschutzverfahren durchführen zu müssen.“

Fazit

- ▶ Berechtigtes Interesse an KI-Training möglich
- ▶ Aber: zahlreiche Detailfragen noch zu klären (insb. DS-Infos und Zweckkompatibilität)
- ▶ Auch EDSA lässt diese offen

Ausblick

- ▶ „AI Data Protection Regulation“ (Universität Wien - Wissenschaftlicher Diskussionsentwurf)
- ▶ Richtlinie (EU) 2019/790 (DSM-RL) Art. 4: „Ausnahmen und Beschränkungen für das Text und Data Mining“ (§ 44b UrhG) analog?
- ▶ Next Stop EuGH: weitere EuGH-Entscheidungen sind zu erwarten / abzuwarten

Vielen Dank!
Fragen?
Diskussion!