

Training generativer KI-Modelle mit urheberrechtlich geschützten Inhalten

Ekaterina Filikhina, LL.M.

DLA Piper UK LLP

Herbstakademie 2025

Inhaltsverzeichnis

1. Zulässigkeit des Trainings generativer KI mit urheberrechtlich geschützten Inhalten
2. Die Grundprinzipien des Trainings generativer KI-Modelle
3. Schutzgegenstand und urheberrechtsrelevante Verletzungshandlung
4. Die Schranke des § 44b UrhG, Art. 4 DSM-RL und der Nutzungsvorbehalt
5. Sonstige Schranken
6. Gesamtergebnis und Ausblick

1. Zulässigkeit des Trainings generativer KI mit urheberrechtlich geschützten Inhalten

Zulässigkeit des Trainings generativer KI mit urheberrechtlich geschützten Inhalten

- ▶ Vorwurf: KI-Konzerne würden sich durch das TDM kostenfrei das weltweite Wissen einverleiben und hierdurch die Kreativbranche gefährden
- ▶ Rechtsfragen bereits seit Jahren Gegenstand einer juristischen Diskussion und diversen Klagen
 - ▶ z.B. Getty Images (US), Inc. gegen Stability AI, Inc. in den USA
 - ▶ LG Hamburg Urteil vom 27. September 2024 (nicht rechtskräftig) in Deutschland
- ▶ Ist das Training generativer KI mit urheberrechtlich geschützten Inhalten in Deutschland zulässig?

2. Die Grundprinzipien des Trainings generativer KI-Modelle

Die Grundprinzipien des Trainings generativer KI-Modelle

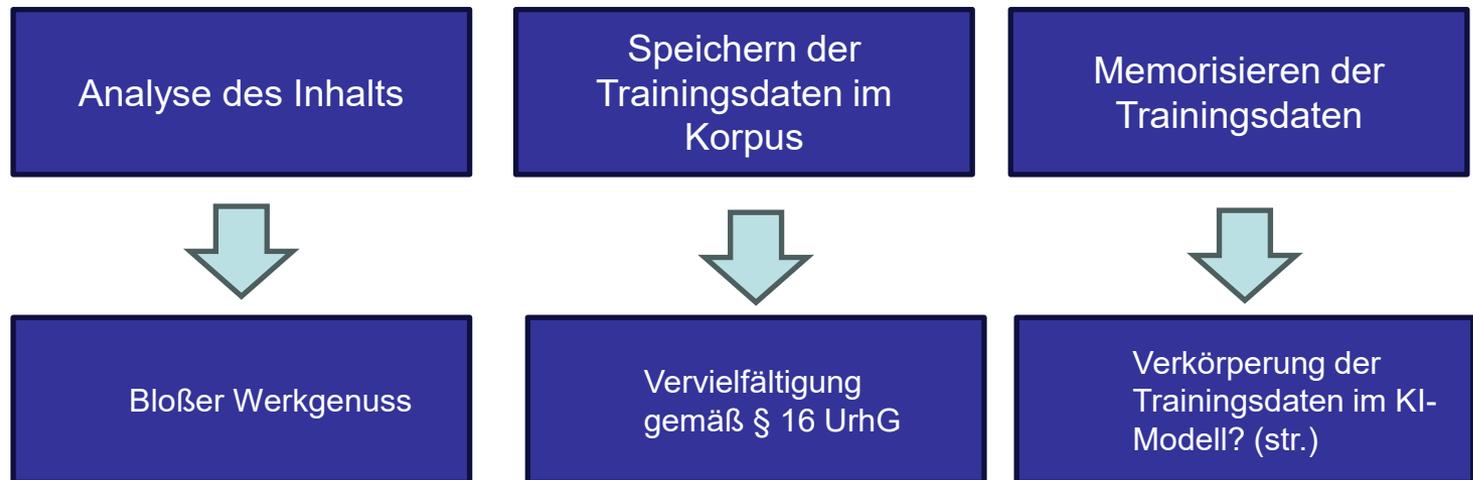
- ▶ Das Trainieren von KI-Modellen erfordert große Mengen an Trainingsdatensätzen.
- ▶ Material wird durch KI-Unternehmen mittels Webcrawler teils aus frei zugänglichen Quellen gesammelt.
- ▶ Ausgangspunkt ist das Entnehmen der Daten aus der Ausgangsquelle und der anschließenden Aufbereitung dieser Daten.
- ▶ Das Ergebnis der Aufbereitung der Daten ist der sog. Korpus. Dieser wird nach Abschluss der Aufbereitungsphase im Rahmen einer Analysephase automatisiert untersucht.
- ▶ TDM ist die automatisierte Analyse von einzelnen oder mehreren digitalen oder digitalisierten Werken, um daraus Informationen insbesondere über Muster, Trends und Korrelationen zu gewinnen.



3. Schutzgegenstand und urheberrechtsrelevante Verletzungshandlung

Schutzgegenstand und urheberrechtsrelevante Verletzungshandlung

- ▶ Schutzgegenstand: an den gesammelten Trainingsdaten bestehen in der Regel Urheber- bzw. Leistungsschutzrechte
- ▶ Verletzungshandlung:



4. Die Schranke des § 44b UrhG, Art. 4 DSM-RL und der Nutzungsvorbehalt

Die Schranke des § 44b UrhG, Art. 4 DSM-RL und der Nutzungsvorbehalt

Gesetz über Urheberrecht und verwandte Schutzrechte (Urheberrechtsgesetz) § 44b Text und Data Mining

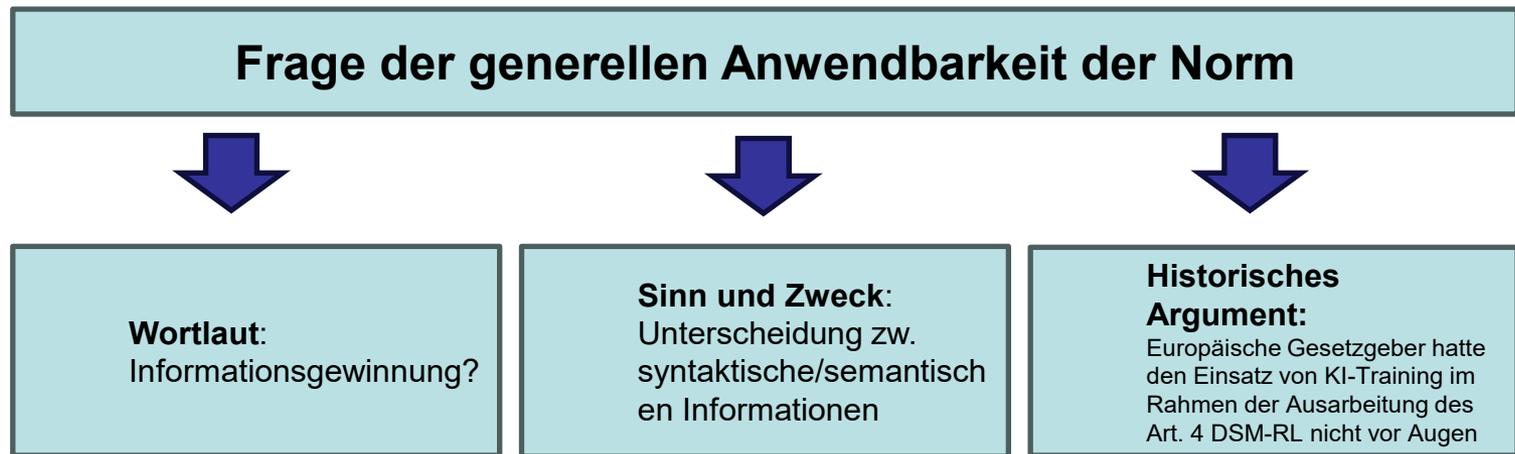
(1) Text und Data Mining ist die automatisierte Analyse von einzelnen oder mehreren digitalen oder digitalisierten Werken, um daraus Informationen insbesondere über Muster, Trends und Korrelationen zu gewinnen.

(2) Zulässig sind Vervielfältigungen von rechtmäßig zugänglichen Werken für das Text und Data Mining. Die Vervielfältigungen sind zu löschen, wenn sie für das Text und Data Mining nicht mehr erforderlich sind.

(3) Nutzungen nach Absatz 2 Satz 1 sind nur zulässig, wenn der Rechtsinhaber sich diese nicht vorbehalten hat. Ein Nutzungsvorbehalt bei online zugänglichen Werken ist nur dann wirksam, wenn er in maschinenlesbarer Form erfolgt.

- ▶ Im Rahmen des KI-Trainings finden Vervielfältigungshandlungen statt, die durch Schrankenregelungen gedeckt sein könnten.
- ▶ Nach § 44b Abs. 2 Satz 1 UrhG sind Vervielfältigungen von urheberrechtlich geschützten Werken für das KI-Training zulässig.
- ▶ TDM-Schranke des § 44b UrhG wurde im Rahmen der Umsetzung der DSM-RL weitgehend inhaltsgleich in das UrhG aufgenommen

Die Schranke des § 44b UrhG, Art. 4 DSM-RL und der Nutzungsvorbehalt



Streit aufgrund von Art. 53 Abs. 1 Buchst. c KI-VO
nun obsolet?

Die Schranke des § 44b UrhG, Art. 4 DSM-RL und der Nutzungsvorbehalt

- ▶ Sofern online zugängliche Werke betroffen sind, setzt der Wortlaut des § 44b Abs. 3 Satz 1 UrhG voraus, dass der Nutzungsvorbehalt in maschinenlesbarer Form erklärt werden muss.
- ▶ Begriff der „Maschinenlesbarkeit“
 - ▶ LG Hamburg Urteil vom 27. September 2024 (nicht rechtskräftig): in natürlicher Sprache verfasste und auf der Webseite der Beklagten veröffentlichte Nutzungsvorbehalt genügt
 - ▶ a.A. technischen Codierung erforderlich
- ▶ Art. 53 Abs. 1 Buchst. c KI-VO: keine konkreten Vorgaben hinsichtlich der Auslegung dieses Merkmals
- ▶ Auslegung des LG Hamburg widerspricht dem Ziel des § 44b Abs. 3 UrhG ein effizientes und rechtssicheres Erkennen eines Nutzungsvorbehalts zu ermöglichen
- ▶ Lösungsvorschläge:
 - ▶ Einrichtung eines Registers für die Erklärung des Nutzungsvorbehalts
 - ▶ Praxisleitfaden: zweite Entwurf des KI-Verhaltenskodex beinhaltet Verweis auf das Robot Exclusion Protocol

5. Sonstige Schranken

Sonstige Schranken

Schranke des § 44a UrhG, Art. 5 Abs. 1 InfoSoc-RL

Gesetz über Urheberrecht und verwandte Schutzrechte (Urheberrechtsgesetz) § 44a Vorübergehende Vervielfältigungshandlungen

Zulässig sind vorübergehende Vervielfältigungshandlungen, die flüchtig oder begleitend sind und einen integralen und wesentlichen Teil eines technischen Verfahrens darstellen und deren alleiniger Zweck es ist,

1. eine Übertragung in einem Netz zwischen Dritten durch einen Vermittler oder
2. eine rechtmäßige Nutzung

eines Werkes oder sonstigen Schutzgegenstands zu ermöglichen, und die keine eigenständige wirtschaftliche Bedeutung haben.



Vervielfältigungen im Zusammenhang mit dem KI-
Training bei denen die Trainingsdaten lediglich
zwischengespeichert und unmittelbar danach gelöscht
werden

Schranke des § 60d UrhG, Art. 3 DSM-RL

Gesetz über Urheberrecht und verwandte Schutzrechte (Urheberrechtsgesetz) § 60d Text und Data Mining für Zwecke der wissenschaftlichen Forschung

(1) Vervielfältigungen für Text und Data Mining (§ 44b Absatz 1 und 2 Satz 1) sind für Zwecke der wissenschaftlichen Forschung nach Maßgabe der nachfolgenden Bestimmungen zulässig.

(2) Zu Vervielfältigungen berechtigt sind Forschungsorganisationen. Forschungsorganisationen sind Hochschulen, Forschungsinstitute oder sonstige Einrichtungen, die wissenschaftliche Forschung betreiben, sofern sie

1. nicht kommerzielle Zwecke verfolgen,
2. sämtliche Gewinne in die wissenschaftliche Forschung reinvestieren oder
3. im Rahmen eines staatlich anerkannten Auftrags im öffentlichen Interesse tätig sind.

Nicht nach Satz 1 berechtigt sind Forschungsorganisationen, die mit einem privaten Unternehmen zusammenarbeiten, das einen bestimmenden Einfluss auf die Forschungsorganisation und einen bevorzugten Zugang zu den Ergebnissen der wissenschaftlichen Forschung hat.

(3) Zu Vervielfältigungen berechtigt sind ferner

1. Bibliotheken und Museen, sofern sie öffentlich zugänglich sind, sowie Archive und Einrichtungen im Bereich des Film- oder Tonerbes (Kulturerbe-Einrichtungen),
2. einzelne Forscher, sofern sie nicht kommerzielle Zwecke verfolgen.

(4) Berechtigte nach den Absätzen 2 und 3, die nicht kommerzielle Zwecke verfolgen, dürfen Vervielfältigungen nach Absatz 1 folgenden Personen öffentlich zugänglich machen:



Modifikation der Regelung des § 44b UrhG speziell für
Zwecke der wissenschaftlichen Forschung

6. Gesamtergebnis und Ausblick

Gesamtergebnis und Ausblick

- ▶ Entscheidung des EU-Gesetzgebers zugunsten der Innovationsförderung.
- ▶ Verantwortung der KI-Unternehmen, Nutzungsvorbehalte mit Hilfe modernster Technologien zu erkennen.
- ▶ Die Einzelheiten der technischen Umsetzung befinden sich noch in einem frühen Entwicklungsstadium.
- ▶ Beseitigung der Rechtsunsicherheiten durch
 - ▶ Nutzungsvorbehalte beachten
 - ▶ Abschluss individueller Lizenzverträge mit Rechteinhabern
 - ▶ Abschlusses von kollektiven Lizenzen mit erweiternder Wirkung nach § 51 Abs. 1 VGG
- ▶ Weitere Entwicklungen
 - ▶ 13. November 2024 Klage der GEMA gegen OpenAI
 - ▶ 21. Januar 2025 Klage der GEMA gegen Suno, Inc.
 - ▶ USA: U.S. District Court D. Delaware Opinion v. 11.2.2025 - kommerzielle Nutzung nicht durch Fair Use freigestellt

Vielen Dank für Ihre Aufmerksamkeit!

Info für Rückfragen:

Ekaterina Filikhina, LL.M.

Senior Associate, DLA Piper UK LLP

E-Mail: ekaterina.filikhina@dlapiper.com