

Training generativer KI-Modelle – die rechtliche Perspektive

Mainz Media Forum

9. April 2025

Dr. Lucie Antoine



Agenda

Überblick:
Urheberrecht und
KI-Training



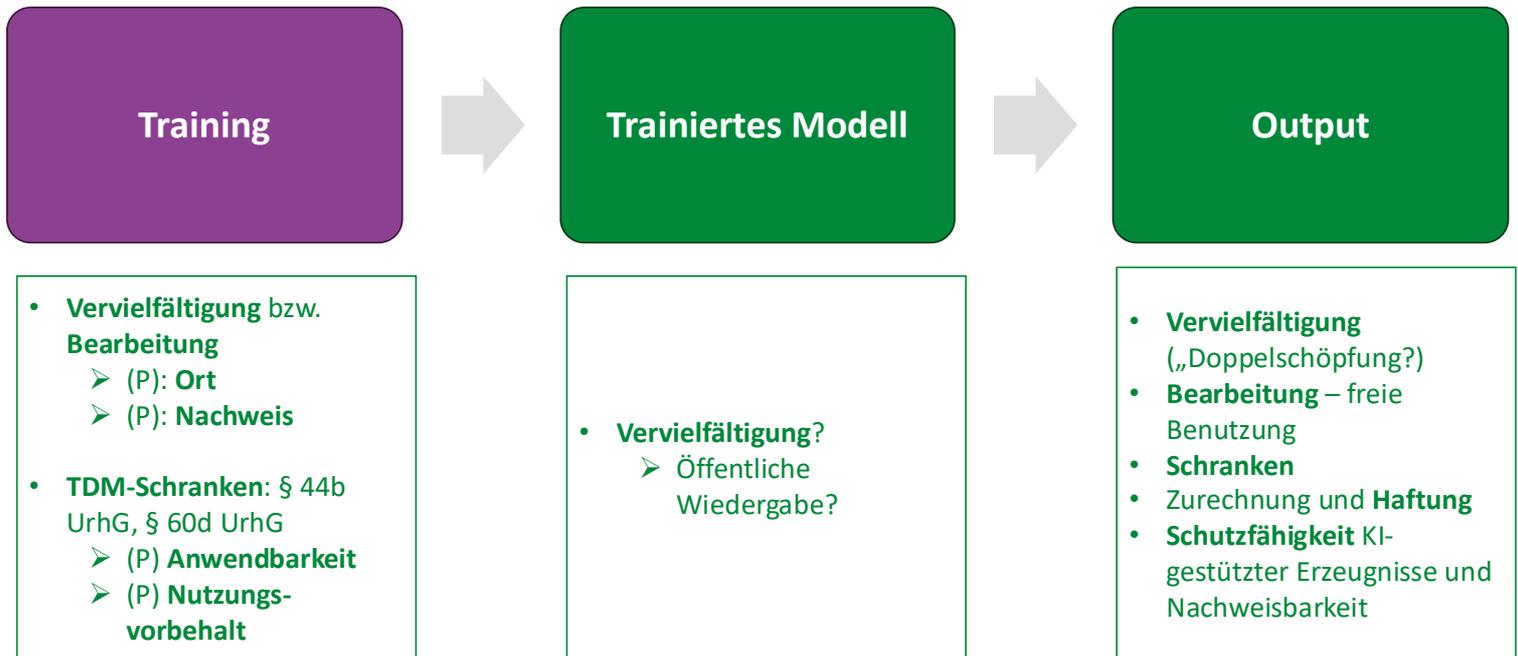
Trainingsvorgang –
urheberrecht-liche
Perspektive



Auswirkungen der
KI-Verordnung



**Ausblick und
Perspektiven**



KI-Training mit urheber- bzw. leistungsschutzrechtlich geschützten Inhalten (I)

- **Ausgangspunkt:** unzweifelhaft **Vervielfältigungshandlungen** (inkl. Bearbeitung)
 - Urheberrechtlich **relevante Nutzung** (§ 16 UrhG)
- **Praktisches Problem:** **Nachweisbarkeit** der Nutzung zu Trainingszwecken
 - **Exemplarisches Beispiel:** GEMA vs SUNO
 - **Ansatz der KI-VO:** Anbieter von KI-Modellen mit allgemeinem Verwendungszweck zu Erstellung und Veröffentlichung einer **Zusammenfassung** der für Training verwendeten Inhalte verpflichtet
 - **Art. 53 Abs. 1 lit. d KI-VO**

Zusammenfassung der für Training verwendeten Inhalte (Art. 53 Abs. 1 lit. d) KI-VO

- **Art. 53 Abs. 1 lit. d KI-VO:** Anbieter von KI-Modellen mit allgemeinem Verwendungszweck ... **erstellen und veröffentlichen eine hinreichend detaillierte Zusammenfassung der für das Training** des KI-Modells mit allgemeinem Verwendungszweck **verwendeten Inhalte** nach einer vom Büro für Künstliche Intelligenz bereitgestellten **Vorlage**.
 - **Erwgr. 107 KI-VO:** Unter gebührender Berücksichtigung der Notwendigkeit, Geschäftsgeheimnisse und vertrauliche Geschäftsinformationen zu schützen, sollte der **Umfang dieser Zusammenfassung allgemein weitreichend und nicht technisch detailliert sein, um** Parteien mit berechtigtem Interesse, einschließlich der **Inhaber von Urheberrechten, die Ausübung und Durchsetzung ihrer Rechte nach dem Unionsrecht zu erleichtern**, beispielsweise indem die **wichtigsten Datenerhebungen oder Datensätze** aufgeführt werden, die beim Training des Modells verwendet wurden, etwa große **private oder öffentliche Datenbanken oder Datenarchive**, und indem eine **beschreibende Erläuterung anderer verwendeter Datenquellen** bereitgestellt wird.

5

KI-Training mit urheber- bzw. leistungsschutzrechtlich geschützten Inhalten (II)

- **Ausgangspunkt:** unzweifelhaft **Vervielfältigungshandlungen** (inkl. Bearbeitung)
- **Rechtliches Problem: internationale Dimension**
 - **Territorial begrenzte Reichweite des Urheberrechts** (sachrechtliche Ebene): **Nutzungshandlung** im Inland oder mit hinreichendem Inlandsbezug
 - **(P) Lokalisierung der Nutzungshandlung:** Serverstandort, Sitz des Anbieters
 - **Jedenfalls regelmäßig nicht EU – keine inländische Benutzungshandlung**
 - Relevanz auch für **Internationale Zuständigkeit**
 - Anwendbares Recht: **Schutzlandprinzip, Art. 8 Rom II-VO (Verletzungshandlung)**
 - Ansatz der **KI-VO:** Anbieter von KI-Modellen mit allgemeinem Verwendungszweck verpflichtet, **Strategie zur Einhaltung des Urheberrechts der Union** zu implementieren
 - Art. 53 Abs. 1 lit.c KI-VO – aber: **Reichweite?**

6

Die Text-and-Data-Mining-Schranke(n)?

§ 44b UrhG – 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.

§ 60d UrhG – 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.

Freistellung durch die Text-and-Data-Mining-Schranke des § 44b UrhG? (I)

- **Ausgangspunkt: Art. 4 DSM-RL**
- **Anwendbarkeit der TDM-Schranken auf KI-Training?**
 - **Wortlaut:** Definition Text-and-Data-Mining (Muster, Trends, Korrelationen)
 - **Telos:** Innovationsförderung, Lizenzmärkte (Nutzungsvorbehalt)
 - **KI-VO:** Einheit der Rechtsordnung
- **Reichweite: Vervielfältigungen von rechtmäßig zugänglichen Werken**
 - **Vervielfältigung** gem. **Art. 2 InfoSoc-RL** = auch Bearbeitungen (vgl. EuGH Painer, Pelham)
 - **Rechtmäßig zugängliche Werke:** insb. nicht hinter pay wall etc.
 - **Löschpflicht:** wenn Vervielfältigungen nicht mehr erforderlich

Freistellung durch die Text-and-Data-Mining-Schranke des § 44b UrhG? (II)

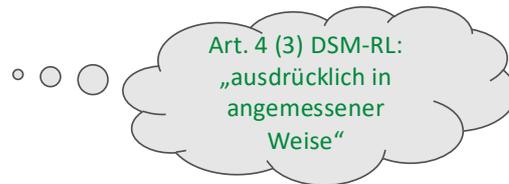
- **Entscheidende Voraussetzung: kein Nutzungsvorbehalt iSd § 44b Abs. 3 UrhG!**

- **Berechtigter (Wer?)**

- **Wortlaut: „Rechtsinhaber“**
- Inhaber einfacher Nutzungsrechte? Plattformen/Intermediäre? „Jeder“?

- **Ausdrücklichkeit (Was?)**

- Nicht konkludent und hinreichend konkret
- Präzise, aber wie spezifisch/granular?



- **In angemessener Weise (Wie?): insb. Maschinenlesbarkeit**

- Natürliche Sprache? Machine code? Technischer Standard (robots.txt uÄm)?
- „Maschinenverständlichkeit“

LG Hamburg v. 27.9.2024 – 310 O 227/23, GRUR 2024, 1710 – LAION

LAION

Large-scale Artificial Intelligence Open Network

TRULY OPEN AI. 100% NON-PROFIT. 100% FREE.

LAION, as a non-profit organization, provides datasets, tools and models to liberate machine learning research. By doing so, we encourage open public education and a more environment-friendly use of resources by reusing existing datasets and models.

Re-LAION 5B release (30.08.2024)

LAION-400M

An open dataset containing 400 million English image-text pairs.

LAION-5B

A dataset consisting of 5.85 billion multilingual CLIP-filtered image-text pairs.

Clip H/14

The largest CLIP (Contrastive Language-Image Pre-training) vision transformer model.

LAION-Aesthetics

A subset of LAION-5B filtered by a model trained to score aesthetically pleasing images.

LG Hamburg v. 27.9.2024 – 310 O 227/23, GRUR 2024, 1710 – LAION

- **§ 44b UrhG: Wirksamer Nutzungsvorbehalt iSd § 44b Abs. 3 UrhG – obiter dictum**
 - **Berechtigter: Urheber, Rechtsnachfolger oder Inhaber abgeleiteter Nutzungsrechte**
 - **Urheber** kann sich in der Folge **auf** entsprechende **Erklärung berufen**
 - **Ausdrückliche Erklärung: *expressis verbis* (nicht konkludent) und *zielgenau* (konkret-individuell)**
 - Erforderlich: **Bestimmbarkeit des Inhalts und der Nutzungshandlung**
 - **Deshalb:** Nutzungsvorbehalt für **sämtliche auf Website** eingestellte Bilder möglich
 - **Maschinenlesbarkeit: = Maschinenverständlichkeit**
 - **Sinn und Zweck:** automatisierte Abfrage durch Crawler
 - **Natürliche Sprache kann grds. maschinenverständlich sein**
 - **Entscheidend: state of the art** (mVerw. auf Art. 53 I lit. KI-VO „modernste Technologien“)
 - **Telos:** angemessener **Interessenausgleich**
 - **Rechtsinhaber:** Nutzungsvorbehalt ist lediglich in „angemessener“ Weise zu erklären
 - **Nutzer-/Anbieterseite:** Anwendung entsprechender technischer Werkzeuge **zumutbar**

11

ECLI:NL:RBAMS:2024:6563

Uitspraak delen

Instantie	Rechtbank Amsterdam
Datum uitspraak	30-10-2024
Datum publicatie	13-11-2024
Zaaknummer	C/13/737170 / HA ZA 23-690
Rechtsgebieden	Intellectueel-eigendomsrecht
Bijzondere kenmerken	Bodemzaak
Inhoudsindicatie	De gedaagde partij heeft als bedrijfsvoering het maken van persoverzichten. De vraag is of gedaagde daarmee een inbreuk maakt op het persuitgeversrecht, dan wel het auteursrecht of het databankenrecht van de uitgevers van diverse kranten en bladen. De rechtbank oordeelt van niet.
Vindplaatsen	Rechtspraak.nl Mediaforum 2025-1, nr. 2 met annotatie van A.B.E. dos Santos Gil IER 2025/7 met annotatie van J.M.B. Selgnette
	Verrijkte uitspraak

[Urteil \(NL\)](#)

[Dazu: Cerri, IPKat v. 16.2.2025](#)

DR. LUCIE ANTOINE

NL

HU

Mezei, The Multi-layered Regulation of Rights Reservation (Opt-out) Under EU Copyright Law and the AI Act, 2025, S. 16 f. ([SSRN](#))

⁹⁷ Municipal Court of Appeals, Case 9.Pf.20.353/2024/6-II, 3 December 2024, para. [51]. (Hungarian court decisions might be located via <https://eakta.birosag.hu/anonimizalt-hatarozatok> in anonymized forms. This actual court ruling might be found by using the keyword 20.353/2024.)

In its 3 December 2024 judgment, the Municipal Court of Appeals reversed the trial court’s refusal of such defence, and concluded that “[i]t was not disputed that the defendant had lawful access to the plaintiff’s press publications and did not circumvent the technical measures taken by the plaintiff to access them. Moreover, the robots exclusion protocol assigned to the plaintiff’s website allowed all search robots, including the defendant’s search robot, to crawl and index, and the plaintiff did not object to the indexing in the form required by law. The robots exclusion protocol is not of significance for consent, but primarily for the fact that it is – in accordance with the provisions of the Copyright Act – a machine-readable form by which the right holder could object to the text- and data-mining by the search engine. Since the defendant does not save the download of the page required for crawling, as it does not keep a copy of the page, and the data and information contained in the index are not copies, the statutory condition in Article 35/A(1)(c) [of the Copyright Act] is also satisfied”.⁹⁷

12

LAION weitergedacht ... (I)

- **Weiter Kreis der zur Erklärung des Opt outs Berechtigten:**
 - Im Ausgangspunkt **überzeugend** (Anknüpfungspunkt §§ 164 ff. BGB)
- **Großzügiges Verständnis der Maschinenlesbarkeit iSe Maschinenverständlichkeit**
 - **Einzelfallbetrachtung: Auffindbarkeit** als gleichermaßen entscheidender Faktor
 - Überdies: **Verhältnismäßigkeit** des Aufwands (vgl. Sinn und Zweck automatisierten Auslesens)
 - Verteilung der **Verantwortungsbereiche überzeugend**
- **Praktische Folge: Großzügiger Maßstab hins. Berechtigten und Maschinenlesbarkeit resultiert in „mehr“ Nutzungsvorbehalten**
 - (P): **Zuordnung & „Mitwandern“ – ggf. Widersprüchlichkeit – Prüfaufwand**

13

LAION weitergedacht ... (II)

- **Ausgangspunkt: Beweislastverteilung (EU + D)**
 - **Nutzer** (KI-Anbieter): **Kein Vorliegen** eines Nutzungsvorbehalts
 - **Rechtsinhaber: Maschinenlesbarkeit**
- **Praktischer Ansatzpunkt: Registerlösungen (EUIPO uÄm)?**
- **Rechtlicher Ansatzpunkt: Doch engere Auslegung der Anforderungen an Nutzungsvorbehalt?**
- **Technischer Ansatzpunkt: Kennzeichnung der Ursprungsquelle? Rolle neuer technischer Standards?**
- **Bei großzügigem Maßstab hins. Berechtigten und Maschinenlesbarkeit:**
 - **Erweiterte Prüfpflichten** der Nutzer (KI-Anbieter) hins. Nicht-Vorliegen eines Opt out (**Abgleich**)?
 - Rolle von **Plattformen**?
 - Besondere („vertrauenswürdige“) **Stellen** für Opt out?

14

Relevante Regelungen der KI-VO für das Urheberrecht

Verbotene Praktiken [inakzeptables Risiko] (Kapitel II)

Hochrisiko-KI-Systeme (Kapitel III)

Transparenzpflichten für bestimmte KI-Systeme [beschränktes Risiko] (Kapitel IV)

Art. 50 Abs. 2 KI-VO: Kennzeichnungspflicht für KI-generierten Output

KI-Modelle mit allgemeinem Verwendungszweck (Kapitel V)

Art. 53 Abs. 1 lit. c KI-VO: Strategie zur Einhaltung des Urheberrechts der Union

Art. 53 Abs. 1 lit. d KI-VO: Zusammenfassung der für Training verwendeten Inhalte

Strategie zur Einhaltung des Urheberrechts der Union (Art. 53 Abs. 1 lit. c) KI-VO

- **Art. 53 Abs. 1 lit. c KI-VO:** Anbieter von KI-Modellen mit allgemeinem Verwendungszweck ... bringen eine Strategie zur Einhaltung des Urheberrechts der Union und damit zusammenhängender Rechte und insbesondere zur Ermittlung und Einhaltung eines gemäß Artikel 4 Absatz 3 der Richtlinie (EU) 2019/790 geltend gemachten Rechtsvorbehalts, auch durch modernste Technologien, auf den Weg;
 - **Erwgr. 106 KI-VO:** (...) Jeder Anbieter, der ein KI-Modell mit allgemeinem Verwendungszweck in der Union in Verkehr bringt, sollte diese Pflicht erfüllen, unabhängig davon, in welchem Hoheitsgebiet die urheberrechtlich relevanten Handlungen, die dem Training dieser KI-Modelle mit allgemeinem Verwendungszweck zugrunde liegen, stattfinden. (...)

Measure I.2.3. Identify and comply with rights reservations when crawling the World Wide Web

- (1) In order to ensure that Signatories will identify and comply with, including through state-of-the-art technologies, machine-readable reservations of rights expressed pursuant to Article 4(3) of Directive (EU) 2019/790 if they use web-crawlers or have such web-crawlers used on their behalf to crawl, scrape and/or otherwise compile data for the purpose of text and data mining according to Article 2(2) of Directive (EU) 2019/790 and the training of their general-purpose AI models, Signatories will
- (a) employ web-crawlers that read and follow instructions expressed in accordance with the Robot Exclusion Protocol (robots.txt), as specified in the Internet Engineering Task Force (IETF) Request for Comments No. 9309, and any subsequent version of this IETF standard, and
 - (b) make best efforts to identify and comply with other appropriate machine-readable protocols to express rights reservations pursuant to Article 4(3) of Directive (EU) 2019/790, for example through asset-based or location-based metadata, that have either resulted from a cross-industry standard-setting process as referred to in paragraph 3 of this Measure or are state-of-the-art and widely adopted by rightsholders, considering different cultural sectors, and generally agreed through an inclusive process based on bona fide discussions to be facilitated at EU level with the involvement of rightsholders, AI providers and other relevant stakeholders as a more immediate solution, while anticipating the development of cross-industry standards referred in paragraph 3.

Code of Practice ([Dritter Entwurf](#), 11. März 2025)

17

Privatrechtliche Durchsetzung der KI-Verordnung?

➤ Schutzgesetz iSd § 823 II BGB bzw. Marktverhaltensregel iSd § 3a UWG?

- Entscheidend etwa **deliktischer Gerichtsstand** (§ 32 ZPO) bei § 823 II BGB!
- Ausgangspunkt: **KI-VO als öffentlich-rechtliche Regulierung**
 - KI-VO dient Schutz von **übergeordneten Interessen der Allgemeinheit** (Gesundheit, Sicherheit) und des **Grundrechtsstandards**
 - **Kein individueller Schadensersatzanspruch** (anders: DSGVO, DSA...)
 - **Eigenständiger Durchsetzungs- und Überwachungsmechanismus** hins. Pflichten der GPAI-Modell-Anbieter durch **AI Office** (nicht: Mitgliedstaaten)
- **Aber: bisher keine klaren Vorgaben auf Unionsebene**
 - **Weites Verständnis** des EuGH in Dieselfällen (NJW 2023, 1111 – Daimler)...
- **In jedem Falle sorgfältig zu prüfen: tatsächlicher Verstoß – Rechtsfolge** (insb. Schaden)
- **Praktische Perspektive:** in EU ansässige Anbieter vs. Anbieter aus Drittstaaten...

Zusammenfassung und Ausblick

- **Entscheidend: Zusammenspiel von Art. 4 Abs. 3 DSM-RL und Art. 53 Abs. 1 lit. c KI-VO**
 - **Reichweite** der erforderlichen Maßnahmen und deren Wirksamkeit unter **Art. 53 Abs. 1 lit. c KI-VO unklar**
 - **Aber:** durch KI-VO noch größerer **Druck auf** konkrete Anforderungen an **Nutzungsvorbehalt** gem. § 44b Abs. 3 UrhG bzw. Art. 4 Abs. 3 DSM-RL – **v.a. Maschinenlesbarkeit**
 - Zentrale **Stellschraube!** Rolle für zukünftige Ansätze wie **Lizenzlösung** etc.
 - Aber: Umsetzung und Effektivität in der Praxis
 - Derzeit erhebliche **Rechtsunsicherheit**
 - Druck auf Lizenzlösungen...?
 - Abzuwarten: **Praxisleitfäden** gem. Art. 56 AI Act
- **Internationale Dimension: USA – Japan – China ...**

19

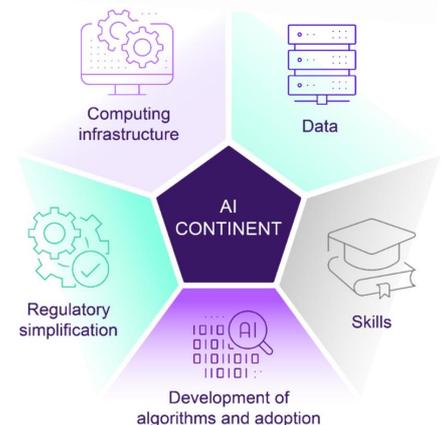
The AI Act's success will primarily depend on how workable its rules are in practice. The current preparatory phase is crucial to achieving a **successful implementation**. The Member States and the Commission, including its AI Office, must step up their efforts to facilitate a smooth and predictable application of the AI Act. As a first step, the Commission is launching the **AI Act Service Desk**, which will be a central information hub on the AI Act, allowing stakeholders to ask for help and receive tailor-made answers. This initiative will provide straightforward and free access to information and guidance on the applicable regulatory framework, which will particularly serve the needs of smaller AI solution providers and deployers. The answers will consist of practical advice that will help to understand and comply with the AI Act. The AI Act Service Desk will be provided by a dedicated team in the AI Office. It will offer an interactive platform where businesses and other stakeholders, including public authorities, will be able to ask questions, get answers and have access to technical tools to help them apply the AI Act, e.g. decision trees and other self-assessment tools.

Brussels, 9.4.2025
COM(2025) 165 final

The **Data Union Strategy** will focus on strengthening the EU's data ecosystem by enhancing interoperability and data availability across sectors, to respond to the scarcity of robust and high-quality data for the training and validation of AI models. It will aim to better align data policies with the needs of businesses, the public sector and society, while fostering a trustworthy environment for data sharing. To achieve this, necessary safeguards will be put in place to ensure the confidentiality, integrity, and security of shared data, thereby promoting a culture of trust and cooperation. Particular attention will be given to streamlining existing data legislation to reduce complexity and administrative burden and to ensure that data governance structures are efficient and effective, based on an inclusive process that takes into account applicable **copyright** legislation.

COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS

AI Continent Action Plan



20

Literatur (Auswahl)

- **Antoine GRUR 2025, 118**
- *de la Durantaye AfP 2024, 9 & ZUM 2025, 165*
- *Dornis GRUR 2024, 1676*
- *Dornis/Stober, Urheberrecht und Training generativer KI-Modelle, 2024*
- *Dregelies GRUR 2024, 1484*
- **Hamann ZGE 2024, 113**
- *Hofmann ZUM 2024, 166*
- *Kuschel ZUM 2025, 174*
- **Leistner GRUR 2024, 1665**
- **Löbbling et al. JIPITEC 2023, 499**
- *Mezei, The Multi-layered Regulation of Rights Reservation (Opt-out) Under EU Copyright Law and the AI Act, 2025 (SSRN)*
- *Nordemann/Rasouli ZUM 2024, 780*
- *Peukert GRUR Int 2024, 497*
- *Quintais, Generative AI, Copyright and the AI Act, 2024 (SSRN)*
- *Rosati, Infringing AI: Liability for AI-generated outputs under international, EU, and UK copyright law, 2024 (SRRN)*
- *Schack NJW 2024, 113 (114)*
- *Sesing-Wagenpfeil ZGE 2024, 212*
- *Stieper/Denga GRUR 2024, 1473*
- *von Welser GRUR-Prax 2023, 516*

Vielen Dank!

Dr. Lucie Antoine

Lucie.Antoine@jura.uni-muenchen.de

