

Gemeinsame Stellungnahme einer breiten Koalition von Urhebern, Künstlern und Rechteinhabern zum 3. Entwurf eines Praxisleitfadens für GPAI-Modelle zur Umsetzung des EU-AI Acts

Die EU-Gesetzgeber haben mit dem AI Act die Bedeutung des geistigen Eigentums für die Fortentwicklung europäischer KI-Modelle anerkannt und deshalb den Entwicklern und Anbietern von KI-Modellen für allgemeine Zwecke in Artikel 53 AI Act Transparenz- und Nachweispflichten auferlegt. Entwickler und Anbieter generativer KI sollen danach verpflichtet werden, konkrete Maßnahmen zur Einhaltung des EU-Urheberrechts zu ergreifen und eine hinreichend detaillierte Zusammenfassung der vervielfältigten und für das Training verwendeten Inhalte bereitzustellen.

Wir merken hierzu – unsere Beiträge vom 20.3.25 zusammenfassend – an:

Der Praxisleitfaden darf nicht im Widerspruch zu geltendem EU-Urheberrecht stehen und muss zusammen mit dem angekündigten Transparenz-Template sicherstellen, dass Urheber, Künstler und Rechteinhaber ihre Rechte gegenüber Anbietern von GPAI-Modellen, prüfen, ausüben und durchsetzen können. Der 3. Entwurf des Praxisleitfadens zur Erstellung einer Urheberrechtsstrategie und der Einhaltung der Nutzungsvorbehalte gemäß Art. 4 Abs. 3 DSM-RL erreicht diese Ziele nach wie vor nicht. Er schafft größere Rechtsunsicherheit, legt das EU-Urheberrecht teilweise falsch aus und untergräbt die Verpflichtungen im AI Act selbst. Die im Praxisleitfaden vorgeschlagenen Maßnahmen, die GPAI-Anbieter ergreifen sollen, sind vage und in ihren Formulierungen selten verpflichtend, so dass sie den Rechteinhabern keine sinnvolle Unterstützung bei der Ausübung oder Durchsetzung ihrer Rechte mehr bieten, sondern diese faktisch sogar verhindern.

Zudem kritisieren wir, dass unsere substanziellen Kommentare zu den vorherigen Entwürfen des Praxisleitfadens, die wir wiederholt, detailliert und konstruktiv schriftlich an das AI Office, die Ministerien und im Rahmen der Stakeholder-Dialoge eingebracht haben, bisher nicht berücksichtigt wurden.

Im Sinne der Schaffung einer hohen Akzeptanz plädieren wir daher dafür, den 3. Entwurf des Praxisleitfadens unter Berücksichtigung der folgenden Aspekte nochmals zu überarbeiten.

1. Rechtskonformität

Anbieter von GPAI-Modellen müssen gemäß Art. 53 Abs. 1c AI Act eine Strategie zur Einhaltung des Urheberrechts und des Nutzungsvorbehalts gemäß Art. 4 Abs. 3 DSM-Richtlinie entwickeln. Die KI-Unternehmen sollen hierbei durch die Ausarbeitung von Praxisleitfäden unterstützt werden (Art. 56 AI Act). Gemäß Art. 53 Abs. 4 AI Act können sich die Anbieter von GPAI-Modellen auf die gemäß Art. 56 AI Act erstellten Praxisleitfäden stützen, um die Einhaltung der Pflichten gemäß Art. 53 Abs. 1 AI Act nachzuweisen. Auch wenn rein rechtlich die Praxisleitfäden – anders als harmonisierte europäische Normen – keine Vermutung ihrer Konformität mit bestehendem Recht darstellen und zudem das Urheberrecht ausdrücklich unberührt lassen sollen, so stellen sie doch für die Anbieter die entscheidende Orientierung für rechtmäßiges Verhalten dar. Deshalb ist es wichtig, dass der Praxisleitfaden zu Art. 53 Abs. 1c AI Act keine Maßnahmen enthält, die suggerieren, dass mit ihrer Befolgung das geltende Urheberrecht eingehalten werde und kein Haftungsrisiko bestünde.

Vielmehr sollten in einem funktionalen Praxisleitfaden urheberrechtlich ungeklärte Fragen explizit benannt werden. Mit der vorliegenden Fassung des Praxisleitfadens werden ungeklärte Auslegungsfragen jedoch zu Lasten der Urheber präjudiziert. Auch wenn es sich beim AI Act um öffentliches Recht handelt, das von Behörden überwacht wird, sollte doch im

Interesse der KI-Unternehmen eine öffentlich-rechtliche Compliance auch bedeuten, dass nicht gleichzeitig gegen Privatrecht verstoßen wird und sich die Unternehmen trotz Einhaltung des Praxisleitfadens des Risikos einer privatrechtlichen Klage aussetzen.

2. Konkreter Änderungsbedarf am Code of Practice

a. Allgemeine Pflicht zur Formulierung konkreter Verpflichtungen:

In allen Fällen, in denen GPAI-Modelle rechtlichen Verpflichtungen unterliegen, müssen diese auch im Praxisleitfaden als solche benannt werden. Es muss deutlich werden, dass nur im Ergebnis rechtskonformes Verhalten auch vor Haftung schützt. Formulierungen wie “best efforts”, die nahelegen, dass ein Bemühen um rechtliche Konformität ausreicht, müssen entsprechend in verpflichtende Formulierungen korrigiert werden.

b. Erwägungsgründe

Gemäß Erwägungsgrund b des Praxisleitfadens soll die Einhaltung der Verpflichtungen im Verhältnis zur Größe und den Kapazitäten der Anbieter stehen; die Interessen von KMU sollen berücksichtigt werden. Art 53 Abs. 1 c AI Act sieht dagegen für Urheberrechtsregeln als zwingendes Gesetzesrecht keine Ausnahmen für KMU vor. Deshalb ist dieser Erwägungsgrund zu streichen.

c. Maßnahme I.2.2 b: Piraterie

Die Formulierung in Maßnahme I.2.2. b des Praxisleitfadens, der regelt, dass Werke auf Piraterie-Domains “möglichst” nicht für Training von KI verwendet werden sollen, sehen wir als besonders problematisch an. Zwingende Voraussetzung für Text und Data Mining ist gemäß Art. 3 und Art. 4 Abs. 1 DSM-Richtlinie der “rechtmäßige Zugang” zu den betreffenden Werken. Dies bedeutet, dass illegal öffentlich zugänglich gemachte Werke nicht verwendet werden dürfen. Auf solchen illegalen Seiten können Rechteinhaber keinen Nutzungsvorbehalt anbringen, dem kommerziellen Text und Data Mining also nicht durch Erklärung widersprechen, ein Recht, das ihnen gesetzlich zusteht. Nach der Logik des Praxisleitfadens würde Art. 4 Abs. 1 Abs. 3 DSM-Richtlinie vollständig leerlaufen. Zumindest bei eindeutig bekannten urheberrechtsverletzenden Webseiten reichen daher unseres Erachtens „angemessene Anstrengungen“ nicht aus.

Wir schlagen folgende Formulierung vor:

„*exclude* from their web-crawling Internet domains that make available to the public copyright-infringing content and have no substantial legitimate uses (“piracy domains”), *in particular* when they are recognized as such by courts or public authorities ...”

Die Diskussion darüber, ob Art. 3 oder 4 DSM-RL nur die Rechtmäßigkeit des Zugriffs oder auch die der Zugänglichmachung regeln, zeigt, dass bei der Formulierung beider Artikel die Auswirkungen auf Rechteinhaber und Urheber durch das Training generativer KI nicht oder jedenfalls nicht ausreichend berücksichtigt wurden.

d. Maßnahme I.2.3.: Kein “Goldstandard” für das robots.txt-Protokoll:

Wir haben mehrfach und für vielfältige Konstellationen dargelegt, dass robots.txt im Zusammenhang mit der Erklärung eines Nutzungsvorbehalts im Sinne von Art. 4 DSM-Richtlinie nicht geeignet ist, eine hervorgehobene Rolle unter den maschinenlesbaren

Nutzungsvorbehalten einzunehmen. Durch die Formulierungen in I.2.3.a. und I.2.3.b wird suggeriert, dass sich ein Unternehmen nur durch Beachtung von robots.txt stets rechtskonform verhält. Das lässt außer Acht, dass es Rechteinhabern nach Art. 4 Abs. 3 DSM-Richtlinie freisteht, andere (oftmals besser geeignete) Lösungen zu nutzen, ohne deshalb eine Missachtung ihres Nutzungsvorbehalts dulden zu müssen. Zwar nimmt Measure I.2.3. in Abs. 2 hierauf Bezug. Das Commitment fokussiert aber allein auf robots.txt.

Zudem ist ein Nutzungsvorbehalt über robots.txt nur dann möglich, wenn Rechteinhaber vor dem Training des Modells wissen, welche Crawler verwendet werden. Maßnahme I.2.4 (4) sieht dagegen noch nicht einmal eine Verpflichtung zur rechtzeitigen Information über die verwendeten Crawler vor.

Der Praxisleitfaden muss für andere Lösungen offenbleiben und jede angemessene und maschinenlesbare Methode der Erklärung eines Nutzungsvorbehalts als gleichwertig anerkennen, wenn er der Rechtslage gerecht werden soll.

Speziell für die Zwecke von Art. 4 DSM-Richtlinie wurde z.B. das »Text and Data Mining Reservation Protocol (TDM-Rep)« entwickelt. Diese Spezifikation des World Wide Web Consortium (W3C) definiert ein einfaches und praktisches Web-Protokoll, das in der Lage ist, den Nutzungsvorbehalt in Bezug auf Text und Data Mining (TDM) für rechtmäßig zugängliche Webinhalte auszudrücken und die Erkennung von TDM-Lizenzierungsrichtlinien für solche Inhalte zu erleichtern. Auch der Internationale Standard ISCC, mit dem sich Rechteerklärungen verbinden lassen, die nicht im Werk eingebettet sind, sondern extern abrufbar sind, werden ebenso negiert, wie weitere existierende internationale Standards für werksgebundene Rechteerklärungen.

e. Maßnahme I.2.4

Wir halten es für nicht ausreichend, wenn sich KI-Unternehmen, die beim Sammeln ihrer Trainingsdaten durch andere Unternehmen crawlen lassen bzw. gecrawlte Inhalte verwenden, sich lediglich um angemessene Informationen “bemühen” müssen. Hier besteht die Gefahr, dass die Verantwortung für die legale Verwendung von gecrawlten Inhalten auf Dritte abgewälzt wird. Zudem wird auch hier wieder nur auf robots.txt verwiesen.

f. Maßnahme I.2.6 (2): Beschwerdemechanismus:

Wir fordern die Streichung des letzten Satzes, der suggeriert, dass KI-Unternehmen vor angeblich “exzessiven” Beschwerden der Rechteinhaber geschützt werden müssten. Hier werden Rechteinhaber unter einen Generalverdacht gestellt, mit der Konsequenz, dass die Regelung eine Ausrede bietet, in Verletzungsfällen berechnete Beschwerden ignorieren zu können. Ein wirksames Schlichtungsverfahren oder andere wirksame Mechanismen zur Vermeidung von Rechtsstreitigkeiten fehlen dagegen.

g. Auffindbarkeit I. 2. 3 (5)

Schließlich sollte der Praxisleitfaden zumindest für die Unterzeichner, die auch eine Suchmaschine betreiben, eine strikte Verpflichtung enthalten, dass das Opt-Out die Auffindbarkeit von Inhalten nicht beeinträchtigt. Dass der Praxisleitfaden hier nur geringe

Anforderungen enthält, vgl. I. 2.3 (5), ist für journalistische und kreative Inhalte fatal, da die Nutzer oft nur über Suchmaschinen auf Medien-Websites gelangen.

3. Transparenz-Template – hinreichend detaillierte Zusammenfassung

Art. 53 Abs. 1 d) AI Act verpflichtet Anbieter von KI-Modellen mit allgemeinem Verwendungszweck, 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 zur erstellen und zu veröffentlichen.

Aus Sicht von Urhebern und Rechteinhabern ist eine solche Zusammenfassung dann hinreichend detailliert, wenn sie werkbezogene Informationen bereithält. Nur wenn nachvollziehbar ist, welche Werke für das Training der KIs verwendet wurden, können Urheber und Rechteinhaber ihre Rechte wahrnehmen und durchsetzen.

Dass eine lückenlose und werksbezogene Zusammenfassung technisch und wirtschaftlich problemlos umsetzbar ist, wurde nicht zuletzt auch von Prof. Dr.-Ing. Sebastian Stober aufgezeigt.

Fazit

Wir benennen diesen konkreten Verbesserungsbedarf in dem Wunsch, weiterhin konstruktiv am Prozess der Entwicklung des Praxisleitfadens mitzuwirken. Es handelt sich hierbei um für uns essenzielle Punkte, die bereits mehrfach übergangen wurden. Wir möchten davor warnen, einen Praxisleitfaden ohne die Berücksichtigung aller genannten Aspekte zu veröffentlichen. Von einer Einbeziehung aller berechtigten Interessen könnte dann keine Rede sein. Gem. Art. 56 Nr. 6, Abs. 2 AI Act könnte ein solches Dokument durch die EU-Kommission mit allgemeiner Gültigkeit versehen werden und dadurch für noch mehr Rechtsunsicherheit sorgen – aber vor allem den bereits entstandenen Schaden für Urheber, Künstler, Rechteinhaber und die gesamte Kultur- und Kreativwirtschaft und die freien Medien unumkehrbar potenzieren.

Wir appellieren an die Bundesregierung, sich entschlossen für ausnahmslose Transparenz, erleichterte Rechtsdurchsetzung und eine menschenzentrierte Technologie-Regulation einzusetzen. Die Auslegung der KI-Grundverordnung ist der essenzielle Baustein, um das erste global wirksame Gesetz zur Handhabung datenbasierter Systeme so zu gestalten, dass es die Innovationskraft der Quellen jeglicher KI-Systeme, insbesondere der Kultur- und Medienwirtschaft, nicht noch weiter und unwiderruflich zerstört.

Berlin, 24. März 2025

