

Generative KI und Autorenrechte

Marc Scheufen, 30.10.2023

Aufsätze, Studien oder allgemeine Werke der Literatur nicht mehr selbst schreiben, das ist im Zeitalter der Künstlichen Intelligenz (KI) längst keine Utopie mehr. Generative KI, zum Beispiel ChatGPT, stellt dabei den Rechtsrahmen vor ungeahnte Herausforderungen. Diese disruptive Änderung macht auch vor dem Urheberrecht nicht halt, wie der Aufschrei vieler Autoren in den Medien deutlich zeigt.

Das Urheberrecht dient üblicherweise dem Schutz vor dem Kopieren von Werken der Literatur, Wissenschaft und Kunst (Eger/Scheufen, 2012). Aus ökonomischer Sicht soll das Urheberrecht eine Balance zwischen Nutzen (Anreizbildung zur Schaffung eines Werkes) und Kosten (Monopolisierung) des urheberrechtlichen Schutzes schaffen. Deshalb werden der Exklusivität des Urhebers so genannte Schranken des Urheberrechts gegenübergestellt (Scheufen, 2020, 221 ff.). In Form eines Schrankenkatalogs definiert das europäische Urheberrecht konkrete Ausnahmen des urheberrechtlichen Schutzes. Das angelsächsische System versucht hingegen mit der flexibleren Fair-Use Doktrin im Einzelfall Nutzen und Kosten des Schutzes abzuwägen. So kommt beispielsweise der Zitatschranke – die erlaubte Übernahme von fremden Inhalten im eigenen Werk mit Kenntlichmachung – in der Wissenschaft aufgrund ihres kumulativen Charakters eine zentrale Rolle zu.

Im Kontext einer generativen KI stellen sich zwei unterschiedliche Problemstellungen und Perspektiven in Hinsicht des Urheberrechtsschutzes:

- **Input-Perspektive:** Eine generative KI verwendet urheberrechtlich geschützte Texte in Form von Trainingsdaten, um zu lernen und Muster zu erkennen.
- **Output-Perspektive:** Eine generative KI kann auf Basis des eigenen Lernens und der erkannten Muster selbstständig neuartige Texte erzeugen.

Zur Output-Perspektive gibt es schon einen breiten Konsens in der Wissenschaft (Scheufen, 2021) und Politik (Deutscher Ethikrat, 2023), dass eine KI nicht Urheber eigener Texte sein kann und sollte. Mit Blick auf die Input-Perspektive ist die Meinungsbildung trotz einiger Stellungnahmen (zum Beispiel Deutscher Kulturrat, 2023; Initiative Urheberrecht, 2023) allerdings noch nicht abgeschlossen. Auch der sich im Triolog befindliche AI Act bleibt hierzu noch unzureichend.

In diesem Zusammenhang könnten urheberrechtlich geschützte Werke als Input unter die Schrankenregelung für Text und Data Mining (§ 44b UrhG) fallen. Dann wäre die Verwendung urheberrechtlich geschützter Texte für das Training einer KI bei automatisierter Verarbeitung erlaubt. Die mithilfe von Webscraping aus dem Internet erlangten urheberrechtlichen Werke könnten also tendenziell diese Schrankenregelung erfüllen. Ob die Schrankenregelung allerdings greift, ist derzeit noch vollkommen offen. Zudem betonen der Deutsche Kulturrat (2023) sowie die Initiative Urheberrecht (2023) in einer Stellungnahme, dass die Verwendung urheberrechtlicher

Werke für KI-Systeme für nicht-wissenschaftliche, sondern kommerzielle Zwecke nicht von dieser Schrankenregelung eingeschlossen sein dürften. Dies würde die Entwicklung generativer KI unter Verwendung urheberrechtlicher Werke zumindest in Europa unmöglich machen. Vor diesem Hintergrund stellt sich die Frage, ob wir eine den AI Act ergänzende Regulierung zum Schutz der Urheber benötigen. Wenn ja: Welche Elemente sollte eine solche Regelung beinhalten?

KI und Autorenrechte

Gerade bei dem wirtschaftlichen Potenzial in der Entwicklung generativer KI, die erst am Anfang steht, sollte eine urheberrechtlich bedingte Marktzutrittsbarriere für europäische Wettbewerber unbedingt verhindert werden. Nichtsdestotrotz bleiben vor allem aufgrund des naturrechtlichen Charakters des Urheberrechts in Europa zusätzliche Regelungen notwendig. Verschiedene Elemente sind dabei zu berücksichtigen:

Opt-in oder Opt-out. Für Urheber, die keine Nutzung ihrer Werke für KI-Trainingszwecke wünschen, sollte eine Opt-out Möglichkeit bestehen. In einem solchen Fall entscheidet sich der Rechteinhaber explizit für den Ausschluss bei der Verwendung seiner Werke zu KI-Trainingszwecken. Im Gegensatz würde eine Opt-in Regelung, bei der ein Urheber seine explizite Einwilligung erklären muss, eine zusätzliche Marktzutrittsbarriere implementieren. Schließlich müsste der KI-Entwickler eine Lizenzierung mit jedem einzelnen Urheber ersuchen, mit prohibitiv hohen Transaktionskosten. In vielen Fällen dürfte das Nicht-Auffinden eines Urhebers ein Opt-in gar unmöglich machen – beispielsweise bei so genannten Orphan Works im Sinne verwaister Werke, bei der der Rechteinhaber selbst nach sorgfältiger Suche nicht festgestellt werden kann. Bei der großen Menge an Daten, die für das Training einer KI benötigt werden, würde ein Opt-in also die Entwicklung generativer KI-Systeme erheblich erschweren oder unwirtschaftlich machen.

Transparenz. Eine Opt-out Option für den Rechteinhaber setzt wiederum die Notwendigkeit von Transparenzpflichten für KI-Entwickler voraus. Schließlich müsste ein Rechteinhaber zunächst darüber

informiert sein, dass sein Werk als Trainingsdatum einer generativen KI Verwendung findet, um die Opt-out Option zu wählen. Hierzu sollte der sich im Trilog befindliche AI Act konkrete Transparenz- und Offenlegungspflichten definieren, die vor allem in Hinsicht auf große Entwickler-Unternehmen ein Gleichgewicht in der Verhandlung mit den Urhebern erlauben sollte. Die gegenwärtige Formulierung einer Zusammenfassung der Ausbildungsdaten im Art. 28 b Abs. 4 AI Act ist hier leider noch unzureichend ausgestaltet, um eine Identifikation einzelner Inhalte zu gewährleisten (Initiative Urheberrecht, 2023). Die Organisation einer Opt-out Regelung könnte über eine der Verwertungsgesellschaften (z. B. VG Wort für Schriftwerke) umgesetzt und kontrolliert werden. Ein kollektiver Opt-out, wie ihn derzeit beispielsweise die GEMA vorbereitet (Deutscher Kulturrat, 2023), ist allerdings nicht zielführend, weil die Entwicklung generativer KI hiermit unmöglich wird. Das gefährdet die Wettbewerbsposition gegenwärtiger und zukünftiger europäischer Marktteilnehmer.

Vergütung. Analog zum Grundgedanken einer Verwertungsgesellschaft könnte zudem über eine Vergütung der Rechteinhaber nachgedacht werden. Aus ökonomischer Sicht soll eine Verwertungsgesellschaft als Intermediär zwischen Urheber und Nutzer in solchen Fällen auftreten, in denen eine bilaterale Lizenzierung zwischen den Parteien erschwert oder unmöglich wird (Tietzel/Weber, 1994). So kann beispielsweise der Besucher einer Bibliothek auch ohne Zustimmung des Rechteinhabers ein Buch kopieren, während über eine Abgabe für z. B. Kopierpapier oder Kopiergeräte eine Vergütung des Urhebers über die Verwertungsgesellschaft (hier: VG Wort) erfolgt. Hierzu meldet der Urheber seine Werke regelmäßig bei der Verwertungsgesellschaft, die im Auftrag die Rechte der Urheber wahrnimmt (Wahrnehmungsvertrag) und über die Abgabeeinnahmen eine Vergütung erwirkt. Für ein Werk, das als Trainingsdatum Verwendung findet, scheint eine solche Vergütung indes aus mehreren Gründen nicht angebracht. Zunächst sieht die Data-Mining-Schranke eine solche Vergütung nicht vor. Zudem basiert die Forderung nach einer Vergütungspflicht (Initiative Urheberrecht, 2023) auf einem fehlenden technischen Verständnis einer generativen KI – zumal eine KI kein Werk umfassend

kopiert, sondern dieses primär zum Lernen und zur Mustererkennung dient. Aus ökonomischer Sicht könnte eine solche Vergütung nicht nur schwer möglich, sondern auch kontraproduktiv sein. Bei der großen Menge an Daten würde die Entwicklung generativer KI unter Umständen nicht nur teuer, sondern könnte auch gänzlich ausbleiben. Das betrifft vor allem kleine und mittelständische Entwickler-Unternehmen, die im Gegensatz zu Google, Microsoft und Co. nicht über ein entsprechendes finanzielles Polster verfügen. Gleichzeitig könnte eine Abgaberegulierung pro Trainingsdatum dazu führen, dass ein KI-Entwickler weniger Trainingsdaten verwendet als technisch sinnvoll wäre und damit schlechtere KI-Systeme entwickelt. Die Wettbewerbsposition europäischer KI-Entwickler könnte sich damit erheblich verschlechtern.

Fazit

Die Forderung einiger Initiativen nach einer konkreten Regelung zur Verwendung von urheberrechtlichen Werken ist legitim und sollte neben einer Opt-out Option explizite und umfassende Transparenz- und Offenlegungspflichten in den AI Act mit einfließen lassen. Auf ergänzende Maßnahmen neben dem AI Act sollte hingegen verzichtet werden, um nicht die Rechtsunsicherheit bei den Unternehmen durch noch weitere Regulierungsvorgaben zu schüren. Schließlich zeigen Studien, dass vor allem rechtliche Hemmnisse viele Unternehmen immer noch davon abhalten Daten zu bewirtschaften und die Potenziale von Data Sharing zu realisieren (Röhl/Scheufen, 2023). Ähnliche Folgen wären auch für die Entwicklung generativer KI-Systeme denkbar und wahrscheinlich. Eine Vergütung der Rechteinhaber oder gar eine Forderung nach der Gründung einer neuen Verwertungsgesellschaft scheinen hingegen weder technisch, rechtlich noch ökonomisch legitimierbar.

Literatur

Deutscher Ethikrat, 2023, Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz,

Stellungnahme, <https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-mensch-und-maschine.pdf> [30.10.2023]

Deutscher Kulturrat, 2023, Künstliche Intelligenz und Urheberrecht, Stellungnahme, <https://www.kulturrat.de/positionen/kuenstliche-intelligenz-und-urheberrecht> [30.10.2023]

Eger, Thomas / Scheufen, Marc, 2012, Das Urheberrecht im Zeitenwandel: Von Gutenberg zum Cyberspace, in: Müller, Christian / Trosky, Frank / Weber, Marion (Hrsg.), Ökonomik als allgemeine Theorie menschlichen Verhaltens, Schriften zu Ordnungsfragen der Wirtschaft, Band 94, Stuttgart, S. 151-179

Initiative Urheberrecht, 2023, Formulierungsvorschläge für den AI Act (Version vom 14. Juni 2023), https://urheber.info/media/pages/diskurs/positions-papier-zu-kunstlicher-intelligenz/7ddcbd4da8-1695132385/230919_iu-formulierungsvorschlaege-fur-den-ai-act-mit-begrundung_endg.pdf [30.10.2023]

Röhl, Klaus-Heiner / Scheufen, Marc, 2023, Hemmnisse beim Data Sharing: Empirie und Handlungsperspektiven, in: Perspektiven der Wirtschaftspolitik, Jg. 24, Nr. 1, S. 129-144

Scheufen, Marc, 2020, Mikroökonomie und Wirtschaftspolitik: Mit einer Einführung in die ökonomische Analyse des Rechts, 2. Auflage, Berlin

Scheufen, Marc, 2021, Urheberrechtliche Fragen der Künstlichen Intelligenz, in: Leupold, Andreas / Wiebe, Andreas / Glossner, Silke (Hrsg.), IT-Recht, 4. Auflage, München, S. 1033-1039

Tietzel, Manfred / Weber, Marion, 1994, Urheberrechte im Zeitalter der Fotokopie. Die Ökonomik von Verwertungsgesellschaften am Beispiel der VG Wort, in: Ott, Claus / Schäfer, Hans-Bernd (Hrsg.), Ökonomische Analyse der rechtlichen Organisation von Innovationen, Tübingen, S. 128-147