
**Hintergrundpaper der Autor:innen und Übersetzer:innen des
deutschsprachigen Buchsektors: Für einen ethischen Umgang mit KI**

***Text-and-Data-Mining-Ausnahme reformieren • Zustimmungspflicht statt
Nutzungsvorbehalt • Transparenzpflicht und Verantwortungszuweisung***

Autor:innen und Übersetzer:innen sind die Quellen des Buchwirtschaftssektors mit seiner Wertschöpfung von 13,5 Milliarden Euro (davon im Jahr 2022: 9,3 Milliarden Euro Verkaufsumsatz) in Deutschland und rund 23,5 Milliarden Euro im europäischen Sektor. Die Arbeit der rund 30.000 bei den Finanzämtern registrierten Texturheber:innen in Deutschland erlaubt es 113.000 Voll- und Teilzeitbeschäftigten in Verlag, Handel usw., einem Erwerbsleben mit Gehalt nachzugehen, außereuropäischen Plattformen und Aggregatoren, lukrative Geschäftsmodelle zu entwickeln, und Bund und Ländern, ihrem Bildungsauftrag nachzukommen. Als zentrale Ressource dieses Teilmarktes der Kreativ- und Kulturwirtschaft partizipieren die Quellen der Buchwertschöpfungskette jedoch nur mit 5 bis 10 Prozent der Erlöse¹.

Wenn Schriftsteller:innen Werke schaffen, gehen sie privatwirtschaftliches Eigenrisiko ein. Sie werden nicht für ihre Arbeitsleistung (Recherche, geschriebene Seiten) oder Arbeitszeit bezahlt. **Einzig die *Werknutzung* löst eine monetäre Beteiligung an der Verwertung ihrer Arbeit aus**, wie etwa Verkauf, Verleih oder Lizenzen. Die vergütungsfreie Nutzung ihrer Arbeit durch Schranken und Ausnahmen im Urheberrechtsgesetz oder durch gesetzlich erlaubte Nutzungen ohne Einwilligung der Urheber:innen mit nur geringer Vergütung, gehen zu Lasten ihres Einkommens und ihrer Persönlichkeitsrechte. Sie schränken damit die Vielfalt der Stimmen ein und beschädigen damit die Grundlage einer demokratischen Kultur.

Keine Künstliche „Intelligenz“ ohne Kreative Intelligenz.

Seit Jahren wird die professionelle Leistung von Schreibenden und Übersetzenden unvergütet und ohne Zustimmung verwertet, um Softwareprodukte der sogenannten Künstlichen „Intelligenz“ zu entwickeln. Weltweit arbeiten Unternehmen wie Oracle, Alibaba, Google, Microsoft, OpenAI, Nvidia oder Apple und Amazon an Textgeneratoren, maschinellen Übersetzungssystemen und synthetischer Vertonung von Textwerken. Die Datensätze für das Training von Übersetzungs- und Selbstschreibsoftware beruhen neben im Internet zugänglichen gemeinfreien Textwerken auf urheberrechtlich geschützten und digitalisierten Büchern der Jahre 2013–2021 und jünger. Die drei Korpora, die für das Learning der maschinellen Sprachprogramme genutzt wurden, Books1, Books2, sowie Books3, stammen zudem teilweise aus illegalen Quellen², wie etwa Bit Torrent-Seiten; der Hersteller des Programms ChatGPT weigert sich, transparente Auskünfte über den Datensatz zu erteilen.

Text-and-Data-Mining-Schranke reformieren, Zustimmungspflicht statt Nutzungsvorbehalt garantieren, Nachweispflicht anpassen

Die seit 7.6.2021 in Kraft getretene Schrankenregelung zum unvergüteten Text-und-Data-Mining macht Unternehmen das Schürfen von Trainingsmaterial aus EU-Ländern allzu leicht – wie etwa aus E-Books, die für Machine Learning der Textgenerator- und Übersetzungssoftware genutzt werden. Urheber:innen werden nicht entschädigt. Gleichzeitig wird ihre Arbeit von diesen Unternehmen

¹ (890 Mio. € bei den Finanzämtern gemeldete Umsätze im Jahr 2019, 792 Mio. € reiner Umsatz im Jahr 2020

² <https://aicopyright.substack.com/p/the-books-used-to-train-llms>

genutzt, um jene Software, die sich aus unseren Köpfen bedient hat, retour nach Europa zu verkaufen und mit hiesigen Textschaffenden zu konkurrieren. **Das für die Datenextraktion von online zugänglichen Inhalten gesetzlich mögliche „maschinenlesbare Opt-out“, der sogenannte Nutzungs- oder Lizenzvorbehalt anstatt einer Zustimmungspflicht, wird bisher nicht angewandt, da es weder eine technische noch vertragliche Ausgestaltung dazu gibt.**

Das Netzwerk Autorenrechte fordert daher:

- Die geltende TDM-Schranke mit vergütungspflichtiger Ausgestaltung verwertungsgesellschaftlich und wie von der VG Wort bereits 2019 gefordert reformieren³.
- Praktikables Opt-out-/Opt-in-Management für jede:n Urheber:in anwendbar und durchsetzbar gestalten.
- Alternativ Lizenzpflichten, die Entwickler:innen von KI-Produkten verpflichten, Kollektiv- oder Individuallizenzen mit Urheber:innen abzuschließen.
- Erlösbeteiligung an Umsätzen aus Schreib- und Übersetzungssoftware für Text-Urheber:innen.
- Nachweis- und Transparenzpflicht über die zum Training missbrauchten Werke muss bei Nutzer:innen und entsprechend Unternehmen liegen.

KI-Kennzeichnungspflicht, Schutz des geistigen Eigentums, der Mensch im Zentrum

Maschinenübersetzung: Presseorgane nutzen maschinelle Übersetzungssysteme, ebenso Fachbuch- und Wissenschaftsverlage. Übersetzer:innen werden immer seltener für das volle Werk engagiert, sondern stattdessen als Ausgleichende/Anpassende, die „Post-Editing“ von maschinell übersetzten Werken vornehmen, die Verantwortung für den Inhalt wird jedoch den menschlichen Übersetzer:innen zugeteilt. Die Arbeitsweise ändert sich, jedoch nicht die Anforderungen an die Ausbildung und die Sensibilität im Umgang mit Texten, zudem wachsen die Sprachvielfalt und die Anzahl der vielsprachigen Texte. Übersetzungsmaschinen kommen auch zum Einsatz, um Honorare zu drücken. Bei maschinell vorübersetzten Texten schwankt die Qualität des Outputs je nach Software stark, dadurch ist der Arbeitsaufwand nicht pauschal und im Vorfeld abzusehen, was zu Selbstausbeutung führt.

Darüber hinaus werden auch verwertungsgesellschaftliche Vergütungsmechanismen tangiert: Ob METIS (Texte im Internet), Bibliothekstantien und Gerätevergütungen für belletristische oder wissenschaftliche Werke, Presseerzeugnisse und vertonter Text und Synchronübersetzungen: Nur menschliche Urheber:innen haben einen Vergütungsanspruch. Bereits jetzt sehen sich Verwertungsgesellschaften veranlasst, KI-Detektoren entwickeln zu lassen, um nicht berechtigte Texte auszusortieren. Die dafür notwendigen Kosten werden von den Geldern entnommen, die Urheber:innen zuständen; damit entsteht ein weiterer fortgesetzter Schaden.

Das Netzwerk Autorenrechte fordert daher:

- Übersetzer:innen entscheiden selbst, ob und inwieweit sie maschinelle Übersetzungssysteme in ihrer Arbeit einsetzen. Bei maschinell vorübersetzten Texten muss für sogenannte „Post-Editing“-Leistungen ein Stundensatz in einer festzulegenden Mindesthöhe vereinbart werden.

³ https://www.vgwort.de/fileadmin/vg-wort/pdf/veroeffentlichungen/stellungnahmen/Stellungnahme_Umsetzung_EU-Richtlinien.pdf

- Verlage haben Autor:innen darüber zu informieren, ob ihre Werke maschinell (teil-)übersetzt sind oder werden.
- Das Urheberrecht muss bei den Übersetzer:innen verbleiben, unabhängig davon, ob eine Maschine zum Einsatz kommt.
- Eine Kennzeichnungspflicht von mittels KI hergestellten Übersetzungen ist einzuführen. Entsprechend sind Regeln für Umfangsgrenzen zu entwickeln.

Textgeneratoren: NLP-Textgeneratoren werden seit Jahren für simulierte Kundendialoge, Kommentator-Bots, Produktbeschreibungen, Gebrauchsanweisungen etc. verwendet; neuere Software-Produkte wie ChatGPT verfassen mehr oder weniger sinnhafte Langtexte. Die Qualität und Verlässlichkeit sind heute bestenfalls suboptimal, doch wird die fortschreitende, rasante Entwicklung bereits in wenigen Wochen „bessere“ Ergebnisse liefern. Hier entstehen Konkurrenzprodukte zu jenen von Autor:innen, die einen massiven Einfluss auf Vergütungsstrukturen und rechtliche Fragen haben werden.

Das Netzwerk Autorenrechte fordert daher:

- Europäischer AI-Act: Bisher ist keine Kennzeichnungspflicht für KI-Kulturwerke vorgesehen, sondern in Artikel 52 (3) sogar eine Ausnahme geschaffen worden. Dies sollte keinesfalls im Trilog und Europäischen Rat Zustimmung finden.
- Eine Kennzeichnungspflicht von mittels KI hergestellten Presse-, Buch-, und Textwerken ist einzuführen.
- Vor Verabschiedung des AI-Acts ist ein Stakeholder-Dialog mit der Kultur- und Kreativwirtschaft, insbesondere den Urheber:innen mit allen Ressorts der Kultur-, Digital-Rechts- und Bildungspolitik, zusammenzusetzen, um Risiken zu bestimmen und Maßnahmen zum Schutz der Urheber:innen, Kulturnutzer:innen und der Wirtschaftskraft zu entwerfen.

Synthetische Hörbuchvertonung: Apple und GooglePlay vertonen Textwerke und Bücher mit künstlichen Stimmen; einige dieser Stimmen sind bestimmten Personen nachempfunden, so wurde z. B. die Stimme der Schauspieler Emma Watson (Harry Potter-Filme) geklont, und für eine Einlesung von „Mein Kampf“ missbraucht. Das geistige Eigentum der Hörbuchsprecher:innen an ihrer Stimme und das Urheberrecht der Texturheber:innen wird damit umgangen. Im Rahmen des BFGS und der EU-Richtlinie zur Barrierefreiheit sollen E-Books ab 2025 zudem „barrierefrei“ werden. Gerätehersteller bieten bereits integrierte Vorlese-KI an. Wenn der Kreis der BFGS-Berechtigten unkontrolliert bleibt, bedroht dies zusätzlich die Lizenzfreiheit der Texturheber:innen.

Das Netzwerk Autorenrechte fordert daher:

- BFGS / EU-Richtlinie zur Barrierefreiheit: Konsultationen zur Auswirkungen text-to-speech bei Gefahr der Unterlaufung der Lizenzfreiheit der Urheber:innen, und des Leistungsschutzrechtes der Hörbuchsprecher:innen.
- Kennzeichnungspflicht für mittels KI hergestellte Audiobooks.
- Das Lizenzrecht der Texturheber:innen darf nicht unterlaufen werden, auch nicht beim Einsatz von synthetischen Stimmen.
- Beim Einsatz von Klon-Stimmen, die von lebenden Personen kopiert und weiterentwickelt wurden, ist die explizite Zustimmung der Sprecher:innen sowie ihrer Rechtsnachfolger einzuholen, und diese sind angemessen an Erlösen zu beteiligen.

Auswahl und Sichtbarkeit von Buchwerken: Algorithmen kuratieren seit Jahren Empfehlungen im digitalen Umfeld, basierend auf kommerziellem Erfolg. Software „prüft“ unveröffentlichte Manuskripte auf „Bestsellerpotenzial“. Dies hat destruktive Auswirkungen auf Zugang, auf Rezeption, auf die Wertschätzung und die inhaltliche Vielfalt.

Das Netzwerk Autorenrechte fordert daher:

- Verlage und Distributoren haben Auskunft über maschinelle Kuratierung zu erteilen.

Verantwortung und Ethik:

Die Qualität jeder generativen sogenannten künstlichen „Intelligenz“ – die nichts mit realer Intelligenz zu tun hat – wird durch die Masse ihres Training-Datensatzes bestimmt. Je mehr Texte, Bilder oder Musik eine KI „lernt“, desto mehr originäre Stile kann sie nachbilden und menschliche Werke nachahmen. Die von KI-Unternehmen verkauften Produkte sind das Ergebnis von der oft illegalen Nutzung Millionen urheberrechtlich geschützten Bildern, Privatfotos, Texten und Leistungen menschlicher Urheber:innen. Gleichzeitig ist die Verantwortungszuweisung bei z. B. Persönlichkeitsrechtsverletzungen oder Desinformation durch maschinell generierte Produkte ungeklärt.

Wir fordern deshalb eine unter ethischen, urheberrechtlichen, wirtschaftlichen, menschenrechtlichen und persönlichkeitsrechtlichen Aspekten weitreichende Regulierung von KI-Systemen inkl. Transparenz und Dokumentation von Daten- und Vorlagen-Nutzungen.

Diese Ausbeutung unserer Arbeit und unserer kreativen, schöpferischen Intelligenz verstößt gegen die Menschenrechte. Sie schadet den Quellen des gesamten Buch- und Kunstmarkts und wird auch durch Kompensationszahlungen z. B. aus TDM nicht annähernd aufgefangen.

Wir stehen in der Morgendämmerung einer Krise, die weit über die Urheber:innen hinaus reicht: Sie wird alle Arten von Arbeitsplätzen und Berufen, von Ausbildung und Studium, von Wissenschaft und Medien betreffen. Die Regeln jedoch sind menschengemacht: und wir fordern, dass sie endlich gemacht werden.

Berlin, 17. April 2023. *Die Unterzeichnenden:*

42er Autoren e.V.

A*dS Autorinnen und Autoren der Schweiz

Autorinnenvereinigung e.V.

Bundeskongress Kinderbuch

Bundesverband junger Autorinnen und Autoren (BVjA)

DELIA

HOMER e.V.

IG Autorinnen Autoren (Österreich)

IG Übersetzerinnen Übersetzer (Österreich)

Mörderische Schwestern e.V.

PAN-Netzwerk Phantastik-Autor:innen

PEN-Zentrum Deutschland

PEN-Zentrum deutschsprachiger Autoren im Ausland

Selfpublisher-Verband e.V.

Das Syndikat e.V.

Verband deutschsprachiger Übersetzerinnen und Übersetzer (VdÜ)

Verband deutscher Schriftsteller*innen (VS in Verdi)

Das Netzwerk Autorenrechte (www.netzwerk-autorenrechte.de) repräsentiert 16 Verbände und über 16.500 Autor:innen und Übersetzer:innen aus Deutschland, Österreich und der Schweiz. Mitglieder: 42erAutoren, A*dS Autorinnen und Autoren der Schweiz, Autorinnenvereinigung e.V., Bundesverband junger Autoren und Autorinnen (BVjA), Bundeskongress Kinderbuch, DELIA, HOMER, IG Autorinnen Autoren, Mörderische Schwestern e.V., Phantastik-Autoren Netzwerk (PAN) e.V., PEN-Zentrum Deutschland, PEN-Zentrum deutschsprachiger Autoren im Ausland, Selfpublisher-Verband e.V., SYNDIKAT – Verein für deutschsprachige Kriminalliteratur, Verband deutschsprachiger Übersetzerinnen und Übersetzer literarischer und wissenschaftlicher Werke e.V. (VdÜ), Verband deutscher Schriftstellerinnen und Schriftsteller (VS) in ver.di.
Kontakt: info@netzwerk-autorenrechte.de