

Das ist kein Kulturgut. Das ist Schüttgut

Ungefragte und unvergütete Nutzung von Werken der Literatur

geschrieben von Nina George

Wir müssen über das Schweigen der Politik sprechen, über diese zutiefst bestürzende Wortlosigkeit gegenüber der Tatsache, dass erfolgreiche generative »KI«-Systeme – wie die stochastische Repetiermaschine ChatGPT, die visuelle Plagiatsschleuder Midjourney, wie Sora, das Videoprogramm, mit dem sich jeder seine eigene Realität basteln kann –, auf jahrzehntelangen, massiven Urheberrechtsverletzungen beruhen.

Wir reden zum Beispiel von 4 Millionen illegitim genutzten Buchwerken aller Genres, vom Publikumsroman bis zum Fachbuch, die in großen Sprachmodellen zusammen mit Presseartikeln, 51 Millionen Amazonrezensionen und Twitter(X)posts als Grundlage für GPT, Bert, Llama oder Copilot dienen. Wir erleben einen seit über zehn Jahren wachsender Berg an »Trainingssets« mit unkreativen Namen wie »the Pile« oder Books1, Books2, Books3. In den Books-Sets stecken zwei meiner Werke fest; dank einiger Daten-Archäologen konnte ich das eruieren, denn wenig überraschend knickern die Oligopole mit Details, um keine Klagen zu riskieren, so sagte etwa Sam Altmann (Open AI). Immerhin wachsen trotz der Heimlichtuerei die Berge, oder vielmehr: Hügel, an Gerichtsverfahren gegen Microsoft, Meta, Alphabet und Konsorten, wie etwa die Sammelklage der US Authors Guild mit 18 Autoren, darunter George R. R. Martin, Jodi Picoult und John Grisham, gegen Open AI. Stets sind es Schriftsteller, die litigieren; das Schweigen der Verlagsgruppen ist fast so auffällig wie das der politischen Entscheider.

Seitens der KI-Unternehmen findet derweil munter hochprofitable Wertschöpfung auf Basis ungefragter, unvergüteter Nutzung statt, und die derzeitige Software kannibalisiert mit Fake-KI-Büchern und unerlaubten KI-Übersetzungen unseren Markt, ahmt reale Neuerscheinungen nach, bevor diese publiziert sind, – und plagiiert auf Prompt (getippte Befehlseingabe) wie »Schreib im Stil von...« Sentenzen, Figuren, Plots, Absätze und Ausdruck menschlicher Schöpfer.

Wer jetzt reflexartig murmelt »Aber § 44b Urheberrechtsgesetz erlaubt doch ...?«, der sei erinnert, dass die Text- und Data-Mining-(TDM-)Ausnahme, die kurz vor Verabschiedung der EU-Richtlinie 2019/790 in den Text gedübelt wurde, erstens nicht retroaktiv gilt und damit keinerlei Schürfen von geschützten Werken für »KI«-Entwicklung vor dem 7. Juni 2021 legitimiert. Zweitens, dass Text und Data-Mining weder technisch noch rechtlich die weitere Verwendung für das Programmieren von generativer Informatik und das Herstellen von wirtschaftlichen Substituten intellektueller Güter abdeckt. Tech-Unternehmen sollten sich genau überlegen, ob sie auf dem sandigen Grund von 44b ein generatives System bauen. Das dann jene Menschen ersetzt, von denen es sich vorher gratis bedient hat, als »low cost alternative« – gern schon verwendet in Agenturen, Pressehäusern, Hörbuchstudios und den einen oder anderen Buchverlagen. Generierte KI-Cover, KI-Übersetzungen, KI-Stimmen, KI-Illus, KI-Klappentexte aus der Imitationsdose, ach, wie geil ist der Geiz.

Während die deutschen Digital- und Wirtschaftsministerien wenig überraschend vom Silicon Valleylein am Neckar träumen, zeigte sich bei den internationalen Urheberrechtskonferenzen im Frühjahr 2024, dass auch juristische und kulturpolitische Ressorts das nonchalante »Schwamm drüber« dem Tacheles vorziehen. Alle streben eilig in die smarte Zukunft, diskutieren gar eigene Rechte für die Plagiatsprodukte – und rufen uns Spaßverderben zu: »Dann optet doch out, wenn ihr kein TDM wollt!« (Opt-out: Rechtevorbekalt erklären, ein Verbotsschild an das eigene Werk bembeln.) Super Idee! Wenn es nicht eine bizarre Sache wäre, im Gesetz den Rechtevorbekalt so zu verlangen, dass er »maschinenlesbar« sein soll. Leider wurde versäumt mitzuteilen, in welcher Maschinensprache; ein Sätzlein im Impressum reicht nicht, tja, und da schauen wir ratlos, und wissen nicht, ob wir jetzt alle 13,8 Millionen in Europa zirkulierenden Buchtitel verbembeln, oder nur die digitalen, oder nur die seit 2022, wer zahlt die Kosten für diese Notwehr, wer kontrolliert, ob der Opt-out in Metadaten, ISCC-Code oder Kryptogramm respektiert wird? Auch Menschen sollten den Vorbehalt lesen können; es sei an die praktisch veranlagten Buchhändler und Bibliothekare

gedacht, die ein Buch in GPT kippen und zusammenfassen lassen, anstatt selbst zu lesen. Zack, ist es unverlernbar in dem Wiederkäu-Automaten, Opt-out keine Option mehr.

Setzen Unternehmen auf maschinelle Piraterieprodukte statt Menschen, muss klar sein: Dieselben Schutzprivilegien wie für Kulturgüter, verringerter Mehrwertsteuersatz oder Buchpreisbindung, kann es für generierte KI-Kotze nicht geben. Das ist kein Kulturgut, das ist Schüttgut.

Maschinenprodukte müssen menschenlesbare Warnungen erhalten; auch, weil der Leser es verdient, informiert zu entscheiden, für was er sein Geld ausgibt. Es geht um

Vergütungslegitimation gegenüber der VG Wort, es geht um korrekten Umgang mit öffentlichen Geldern bei Preisen, Stipendien oder Verlagsförderung, und dass man sich darauf verlassen kann, keine strunzfaulen Prompter zu würdigen.

Gesamtgesellschaftlich rollt eine Konsequenz auf uns zu, der sich die Politik stellen muss. Mehr Maschinenoutput statt menschlicher Arbeit heißt: geringere KSK-Abgaben und höhere Belastungen für den Bund. Geringere Steuerzahlungen der sonst beauftragten menschlichen Schaffenden.

Sinkende Einzahlungen in Vorsorgeleistungen und höhere Altersarmut. Gedeimütigtes

Handtuchwerfen, Fachkraftverluste – insbesondere bei Übersetzung –, Verlernen von

Kulturtechniken. Keinerlei Anreiz für junge Menschen, einen künstlerischen Beruf zu ergreifen. Mit disruptierten Märkten und mangelnder Durchsetzung von Urheberrechten wird den Folgegenerationen eine unheilbare Ungleichheit hinterlassen.

Wer nun weiter schweigt, erklärt sein Einverständnis mit dem schamlosesten Kulturdiebstahl seit Erfindung von 0 und 1.

Dieser Text ist zuerst erschienen in [Politik & Kultur 5/2024](#).