



Schriftsteller Lukas Bärfuss entdeckte seine Bücher auf einer Piratenwebsite und befürchtet, dass sie zum KI-Training verwendet werden.

Fotos: Urs Jaudas, PD

«Diebstahl» oder «Training»? Der Streit um KI-Daten erreicht die Schweiz

Kreative gegen Techriesen Dürfen Konzerne mit Büchern, Filmen oder journalistischen Artikeln ihre Sprachmodelle frei trainieren? Eine Politikerin will das stark einschränken. Experten fürchten, dass damit KI-Forschung in der Schweiz blockiert wird.

Edith Hollenstein

Preisgekrönte Romane wie «Koala» und «Die Krume Brot» – Tausende Stunden investierte Lukas Bärfuss in seine Werke. Doch nun entdeckte er seine Bücher auf einer russischen Piratenwebsite – und somit wurden sie mit grosser Wahrscheinlichkeit von Techfirmen zum Training ihrer KI-Sprachmodelle benutzt. Bärfuss forderte die Website per Mail auf, den «Diebstahl in Zukunft zu unterlassen», doch: «Ich weiss, das war in den Wind gesprochen.»

Seine Erfahrungen stehen symptomatisch für den grossen Konflikt zwischen Kreativschaffenden und Medienhäusern auf der ganzen Welt auf der einen Seite – und den Techkonzernen auf der anderen. Es geht um die Nutzung von geistigem Eigentum – und die Frage, ob und wie dieses für KI-Modelle genutzt werden darf. Dieser Streit ist nun auch in der Schweiz ausgebrochen.

Motion Gössi will geistiges Eigentum vor KI schützen

Für Firmen wie Open AI, Google, Meta oder Anthropic geht es um Milliarden. Wer möglichst schnell an grosse Mengen guter Inhalte kommt, kann bessere Modelle bauen – und sich im rasant wachsenden Markt entsprechende Anteile sichern. Dies wirft zwei Fragen auf: Dürfen Techfirmen diese Inhalte – Video, Audio, Text, Posts in sozialen Medien – für ihre Modelle nutzen? Und wenn ja, wer soll an der Wertschöpfung beteiligt sein, die daraus entsteht? Gesetzgeber – in den USA wie auch in Europa – müssen Regeln finden, die weder die Kreativwirtschaft noch die KI-Entwicklung behindern.

In der Schweiz hat ein politischer Vorstoss diese Diskussion in Gang gebracht: Die Motion von FDP-Nationalrätin Petra Gösli. Sie verlangt eine Anpassung des Urheberrechtsgesetzes, so dass kreative Werke wie Bücher, Zeitungsartikel oder Bilder besser vor der Nutzung durch KI-Anbieter geschützt sind. Gössi schlägt ein Opt-in-Modell vor: KI-Firmen müssten die explizite Zustimmung der Rechteinhaber einholen, bevor sie deren Werke für das Training ihrer Modelle verwenden dürfen.

Beinahe wäre die Motion in Bern widerstandslos durchgesegelt. In der Herbstsession, die nächste Woche beginnt, hätte nur noch der Nationalrat zustimmen müssen. Nun aber schaltet sich mit der Schweizer Techindustrie eine prominente Gegnerin ein – vor allem das Opt-in-Modell ist ihr ein Dorn im Auge.

«Es hätte für die KI-Forschung dramatische Konsequenzen», sagt Imanol Schlag, Computerwissenschaftler an der ETH Zürich. Er leitet ein Projekt der Swiss AI Initiative, das ein neuartiges Large-Language-Modell entwickelt mit helvetischen Werten: transparent, offen und fair. Demnächst soll das Modell veröffentlicht werden. Doch: «Unser Projekt ist in Gefahr», sagt Schlag. Zusammen mit 90 weiteren Wissenschaftlern – die meisten sind Professoren der ETH Zürich, der EPFL Lausanne und des Swiss National Supercomputing Centre in Lugano – hat er einen Brief an Nationalräte unterzeichnet, der vor der Motion Gössi warnt.

Darin heisst es, dass die Motion praktisch nicht umsetzbar sei, da die Zustimmung aller Rechteinhaber das KI-Training in

der Schweiz verunmögliche und die Motion eine «Insellösung» schaffe. Das schwäche den Forschungs- und Innovationsstandort, da KI-Entwickler die Schweiz umgehen würden. Die Wissenschaftler fordern eine sorgfältige Diskussion, bevor «überhastet reguliert wird».

Ganz anders sehen es die Medienhäuser. Denn wie literarische Werke sind auch journalistische Artikel von den Techfirmen zum KI-Training verwendet worden. «Was die KI-Konzerne machen, ist moderne Piraterie. Sie stehlen uns unsere hart recherchierten Inhalte», sagt Verlegerpräsident Andrea Masüger. Er fordert, dass der Bundesrat «keine Zeit verliert».

Es handle sich um einen «Datenklau in enormem Ausmass», sagt Tanja Messerli, Geschäftsführerin des Schweizer Buchhändler- und Verleger-Verbands. KI-Modelle seien nur deshalb so gut, weil sie mit fundierten Werken trainiert worden seien, denn Bücher lieferten «durchdachte Narrative, fundierte Recherchen und ausgefeilte Stilmittel», im Gegensatz zu Social-Media-Beiträgen.

Verleger fordern «faire Abgeltung für Recherche»

Der Brief der rund 90 besorgten Wissenschaftler führt dazu, dass die zuständige Kommission des Nationalrats diese Woche eine zusätzliche Anhörung durchführen wird.

Hinter den Kulissen machen grosse Techfirmen Druck gegen die Motion. Google, das in der Schweiz unter anderem sein KI-Modell Gemini entwickelt, stellt sich klar dagegen. Eine gesetzliche Opt-in-Regelung sei viel restriktiver und einseitiger als das in der EU etablierte Opt-



«Für das Training unseres Modells war so viel Text nötig, wie wenn jeder Schweizer Bürger zehn Bücher geschrieben hätte.»

Imanol Schlag
Leiter des Projekts Swiss AI Initiative an der ETH Zürich

out-Modell, erklärt ein Sprecher. «Eine solche Opt-in-Regelung in der Schweiz hätte negative Konsequenzen für Innovationen und würde den Wirtschafts- und Forschungsstandort im internationalen Wettbewerb signifikant schwächen.»

Verlegerpräsident Masüger hält dagegen: «Es geht uns nicht darum, das Geschäft der Tech-Firmen zu verbieten, sondern um eine faire Abgeltung für Recherche.» Bei der Ausgestaltung des Gesetzes müsse das zuständige Institut für geistiges Eigentum «selbstverständlich auf die Interessen der Schweiz als Standort für Forschung und Innovation Rücksicht nehmen».

Medienhäuser sind in einer Zwickmühle

ETH-Forscher Imanol Schlag sagt, sein Team habe beim Training die existierenden Opt-out-Regelungen respektiert. Er ist zuversichtlich, dass keine Memorisierung von urheberrechtlich geschützten Werken erfolgt ist. Fürs Training sei zwar so viel Text nötig gewesen, «wie wenn jeder Schweizer Bürger zehn Bücher geschrieben hätte». Doch das seien, so Schlag, ausnahmslos öffentliche Daten gewesen ohne Opt-out und «natürlich keine gestohlenen Inhalte».

Doch nicht nur fürs Training der Sprachmodelle sind Texte nötig, sondern auch beim Generieren von Antworten. Dabei müssen sie auf sogenannte Echtzeitdaten – zum Beispiel journalistische Artikel über aktuelle Ereignisse – zugreifen können, um möglichst treffende Antworten zu liefern. Wenn Medienhäuser wie die NZZ, CH Media, Ringier oder Tamedia, der auch diese Redaktion angehört, einen solchen Zugriff auf ihre Artikel ver-

hindern wollen, können sie dies technisch entsprechend kennzeichnen. Doch das bringt sie in eine Zwickmühle: Ihr Geschäftsmodell zählt mitunter darauf, dass ihre Inhalte in den Treffern von Suchmaschinen und KI-Abfragen vorkommen.

Mögliche Lösung: Ein kollektives Lizenzsystem

Entscheidend wird am Ende vor allem sein, wie die Politik die Motion interpretiert – etwa ob die Zustimmung individuell erfolgen muss oder auch kollektiv ausgestaltet sein kann. Der zuständige Justizminister Beat Jans ist für die Motion Gössi. Er sieht die Möglichkeit, sie pragmatisch umzusetzen.

Am Ende könnte ein sogenanntes kollektives Lizenzsystem die Lösung sein: KI-Firmen könnten Pauschalen einzahlen, die durch eine Verwertungsgesellschaft verteilt werden, wie es heute die Suisa im Musikbereich oder Pro Litteris für literarische und journalistische Werke tut. Autoren und Verlage erhielten anteilig eine Vergütung, wenn ihre Werke für KI-Anwendungen genutzt werden – ähnlich wie beim System mit dem «Kopierrappen» für Zeitungsartikel und Bücher, das schon lange in Kraft ist. Pro Litteris zumindest zeigt sich bereit, an Lösungen mitzuwirken.

Auch wenn die KI-Konzerne womöglich ihre Urheberrechte verletzt haben: Weder von Bärfuss noch von anderen Schweizer Autoren ist bekannt, dass sie geklagt oder Anzeige erstattet haben. Sie hoffen, dass ihr Berufsverband aktiv wird. Doch dieser erachtet dies als erst dann möglich, wenn es gesetzliche Vorgaben gibt, die den Umgang mit KI regeln.