



ILLUSTRATION SIMON TANNER / NZZ

KI-Forscher warnen vor «Insellösung» – zu Unrecht

Eine Motion will die Urheberrechte bei digitalen Inhalten im Zeitalter von KI besser schützen. Wissenschaftler schlagen Alarm – dabei liegt die Zukunft der künstlichen Intelligenz in der Verarbeitung von Qualitätsdaten. Von Barnaby Skinner

Die KI-Forschung schlägt Alarm. In einem eindringlichen Brief, welcher der NZZ vorliegt, wenden sich 92 führende Schweizer Professorinnen und Professoren an die Politik in Bundesbern. Ihr Appell richtet sich gegen eine Motion der FDP-Nationalrätin Petra Gössi, die demnächst im Nationalrat beraten wird. Diese will künftig die Urheberrechte bei digitalen Inhalten im Zeitalter von KI und grossen Sprachmodellen besser schützen. Die Motion verlangt einen Paradigmenwechsel: Statt des heute international etablierten Opt-out-Standards, bei dem Rechteinhaber die Nutzung ihrer Inhalte durch KI-Systeme verbieten können, fordert die Motion ein strenges Opt-in-Regime. Jede Nutzung von Inhalten für das Training von KI muss vorher explizit erlaubt werden.

Für die Forschung schiesst dies weit über das Ziel hinaus, geistiges Eigentum vor KI-Missbrauch zu schützen. In ihrer jetzigen Form würde die Motion den gegenwärtigen Wettbewerbsvorteil der Schweiz in KI-Forschung, KI-Innovation und Digitalwirtschaft in einen Nachteil verkehren, behaupten die Wissenschaftler.

Wichtige Bestimmungen des Urheberrechts sollen für KI-Anwendungen ausgehebelt werden. Konkret geht es um die sogenannte Forschungsschranke. Heute erlaubt der Artikel 24 des Schweizer Urheberrechtsgesetzes Forschenden, geschützte Werke für die wissenschaftliche Analyse zu nutzen. Diese generelle Leseerlaubnis in der digitalen Bibliothek des Internets soll KI-Anwendungen nun entzogen werden. Die Konsequenz, so die Forscher, wäre der «faktische Stillstand» für die Entwicklung von KI-Modellen in der Schweiz. Der administrative Aufwand, bei Millionen von Urhebern weltweit eine Einwilligung einzuholen, sei unmöglich zu bewältigen.

Doch ist die Sorge berechtigt? Tatsächlich existieren heute gewaltige Mengen an Daten, die unter Open-Source-Lizenzen wie Creative Commons frei verfügbar sind. Wikipedia, wissenschaftliche Archive oder Software-Code auf Plattformen wie Github – all dies bleibt eine legale und wertvolle Datenquelle. Auch die Annahme, dass neue hochwertige KI-Entwicklungen vom Absaugen des gesamten Internets abhängen, ist zunehmend über-

Anstatt die eigenen Inhalte als Opfer von Datenkraken zu sehen, müssen Unternehmen diese als Datenschatz betrachten, den es zu schützen und gezielt zu nutzen gilt.

holt. Das Mantra «garbage in, garbage out» hat sich bewahrheitet. Es beschreibt eine fundamentale Wahrheit über jedes datenverarbeitende System, von einem einfachen Taschenrechner bis zur komplexesten künstlichen Intelligenz, wonach die Qualität der Ausgabe eines Systems niemals besser sein kann als die Qualität der Eingabe.

Es zeigt sich immer mehr, dass die Zukunft kleineren, spezialisierten Modellen gehört – sogenannten Small Language Models (SLM). Diese werden mit sorgfältig kuratierten, hochqualitativen Datensätzen für spezifische Aufgaben trainiert. So schlägt beispielsweise ein Sprachmodell, das mit dem gesamten Korpus Schweizer Rechtstexte gefüttert wurde, bei juristischen Fragen jedes Allzweckmodell, das auf dem unstrukturierten Textchaos des Internets basiert. Ein anderes Beispiel: Ein SLM, das mit Millionen von wissenschaftlichen Publikationen, klinischen Studien und Patentschriften trainiert wurde, wird sehr viel effizienter verborgene Querverbindungen zwischen Genen, Proteinen und Krankheiten aufdecken, als dies das neueste Molo-Modell von Open AI kann.

Frei verfügbare Basismodelle

Diese Entwicklung wird durch das Verhalten von Big Tech selbst beschleunigt. Elon Musks soziales Netzwerk X ist ein Paradebeispiel. Der Zugang zu Inhalten auf X wird für externe Firmen immer restriktiver und teurer. Der Grund: Die eigene KI von Elon Musk, Grok, nutzt den exklusiven Zugriff auf Echtzeitdaten als entscheidenden Wettbewerbsvorteil. In einer Welt, in der selbst die grössten Plattformen ihre Gärten einmauern, um ihre Datensätze zu schützen und zu kapitalisieren, erscheint die Forderung, dass Schweizer Medienhäuser ihre Inhalte nicht schutzlos preisgeben sollen, umso verständlicher.

Hier eröffnet sich für Schweizer Unternehmen eine historische Chance. Anstatt die eigenen Inhalte als Opfer von Datenkraken zu sehen, müssen sie diese als Datenschatz betrachten, den es zu schützen und gezielt zu nutzen gilt. Ein Umden-

ken, das auch bei vielen Firmen stattfindet. Anstatt die über Jahrzehnte aufgebauten Archive passiv dem Zugriff Dritter zu überlassen, sollte heute jedes Unternehmen, jede Organisation, von klein bis gross, den Aufbau einer eigenen Informationsarchitektur zur Entwicklung eines auf spezialisierte Aufgaben trainierten SLM prüfen.

Hinzu kommt ein entscheidender ökonomischer Faktor: Die Kosten für die Entwicklung solcher spezialisierter Modelle sind drastisch gesunken. Die Zeiten, in denen das Training einer KI dreistellige Millionenbeträge verschlang, sind für diese neue Generation von Modellen vorbei. Statt bei null anzufangen, können Unternehmen heute auf einem riesigen Ökosystem von frei verfügbaren, leistungsstarken Basismodellen aufbauen.

Der eigentliche Aufwand liegt nicht mehr im energieintensiven Grundtraining, sondern in der gezielten Feinabstimmung – dem Veredeln des Basismodells mit dem eigenen, hochqualitativen Datenschatz. Gekoppelt mit der flexiblen Anmietung von Rechenleistung in der Cloud, bringt dies die Entwicklung von hochleistungsfähiger KI in Reichweite von weit mehr Akteuren als nur den Tech-Giganten. Die Entwicklung einer eigenen, massgeschneiderten KI ist damit keine Frage des Milliardenbudgets mehr, sondern eine strategische Entscheidung, die auch für innovative KMU zugänglich geworden ist.

Ein solches Modell beruht auf einer anderen Philosophie: nicht auf einem gesetzlichen Freibrief für alle, sondern auf gezielten, wertorientierten Partnerschaften. Die Wege in der Schweiz sind kurz. Wenn ein Forschungsteam der ETH mit einem seriösen Anliegen Zugang zu bestimmten Daten benötigt, wird man im Dialog eine Lösung finden, die für beide Seiten einen Mehrwert schafft.

Die Suche nach der Balance

Das Ringen um den Zugang zu Daten wird von den Tech-Giganten mit harten Bandagen geführt. Dabei wiederholt sich ein bekanntes Muster aus den Anfängen der sozialen Netzwerke. Damals gelang es den Plattformen, auf der Basis von fremden Inhalten und dem hehren Ziel des freien Zugangs zu Information milliardenstere Imperien zu errichten, ohne die Urheber je fair am Erfolg zu beteiligen. Doch hochwertige, verlässliche Informationen entstehen nicht von selbst. Sie müssen von Journalisten, Forschern und Kreativen erstellt und finanziert werden. Die Strategie der Konzerne zielt darauf ab, den Nutzen – den freien Rohstoff für ihre KI-Modelle – zu maximieren, während die Kosten der Inhaltserstellung bei den ursprünglichen Urhebern verbleiben. Es ist der Versuch, den bestmöglichen Deal für sich auszuhandeln: der, bei dem der Zugang zu wertvollem Material möglichst wenig kostet, idealerweise nichts.

Genau hier schliesst sich der Kreis zur Motion Gössi. Die Notwendigkeit, der Forschung den Zugang zu urheberrechtlich geschützten Daten zu ermöglichen, ist unbestritten. Sie ist die Voraussetzung für unabhängige Innovation. Doch die berechtigte Sorge der Verlage und Motion-Befürworter ist, dass eine zu pauschale Forschungsausnahme zum Einfallstor für ebenjene Tech-Giganten werden könnte. Es ist das Szenario des Trojanischen Pferdes: Ein amerikanischer Konzern initiiert eine «Forschungskoope-ration» mit einer Schweizer Hochschule, um unter dem Deckmantel der Wissenschaft legal auf geschützte Daten zuzugreifen. Die Früchte dieser Arbeit fliessen anschliessend direkt in die kommerziellen Produkte des Konzerns, ohne dass die ursprünglichen Rechteinhaber je entschädigt werden.

Der Vorstoss von Nationalrätin Gössi trifft damit einen Nerv. Der Grundgedanke, die Rechte von Kreativen und Verlagen im Zeitalter der KI zu stärken, ist nicht nur legitim, sondern notwendig. Nur sind die Folgen einer solchen Regulierung kaum abzuschätzen. Die Geschichte der Technologieregulierung ist voller Beispiele, bei denen gut gemeinte Gesetze unbeabsichtigte Nebenwirkungen hatten. Im Idealfall entsteht ein fairer Markt: Schweizer Medienhäuser und Urheber erhalten eine neue Einnahmequelle durch Lizenzverträge, und die Tech-Konzerne werden gezwungen, für die Nutzung von Qualitätsdaten zu bezahlen. Die Befürchtung, wie sie im Brief der Wissenschaftler zum Ausdruck kommt, ist jedoch, dass das Gegenteil eintritt. Um dem Rechtsrisiko aus dem Weg zu gehen, könnten globale Anbieter ihre Dienste für die Schweiz schlicht sperren. Die Folge wäre eine digitale Selbst-Isolation, der Anschluss an die Weltspitze würde Schweizer Unternehmen verbaut. Die Wahrheit liegt wohl irgendwo dazwischen.

Die entscheidende Aufgabe für die Politik ist es nun, eine intelligente Balance zu finden. Wie kann ein Gesetz aussehen, das den Schutz des geistigen Eigentums garantiert und strategische Partnerschaften von privaten Unternehmen fördert, ohne der öffentlichen Grundlagenforschung – dem Motor für unerwartete Durchbrüche – die Luft abzuschneiden?

Es gilt zu prüfen, wie Ausnahmeregelungen für die nichtkommerzielle Forschung gefunden werden können, um ihre Innovationskraft zu sichern. Es gilt, den richtigen Impuls der Motion aufzunehmen und sie so zu verfeinern, dass sie die Urheberrechte in Zeiten von KI schützt, ohne den Wirtschafts- und Forschungsstandort Schweiz zu schwächen.