Die KI(rche) im Dorf lassen

Plädoyer für eine klima- und sozialverträgliche Nutzung Künstlicher Intelligenz

1. Ein Dimensionssprung und seine Folgen

Künstliche Intelligenz (KI) ist ein weites Feld, es ist viele Jahrzehnte alt und hat erfolgreich zwei sogenannte „KI-Winter" (Phasen reduzierter Förderung) überstanden. Ausgelöst durch die jüngsten Entwicklungen in einem Unterfeld der KI, dem Machine Learning, erfreut es sich nun seit einigen Jahren einer enormen Aufmerksamkeit auch über die Fachwelt hinaus. Die neueste Generation von Methoden (ab ca. 2018) bringt allerdings auch neue Herausforderungen mit sich, schon allein aufgrund ihrer Dimensionen. Im Folgenden verwende ich den Begriff „Large AI" (LAI) und meine damit Machine-Learning-Modelle, die eine gewisse Größe überschreiten, in Bezug auf die benötigten Trainingsdaten, Parameter, Ressourcen für den Betrieb darauf aufsetzender Anwendungen, etc. Dazu gehört insbesondere die generative KI (genKI).[1]

LAI hat einen Preis. Immer. Wenn wir ihn zahlen, merken wir es oft nicht. Den größten Preis zahlen meist nicht wir als Nutzende direkt, sondern andere – Menschen, die sich mit LAI-generiertem Output befassen müssen, Menschen in weniger privilegierten Teilen der Welt, zukünftige Generationen. Entsprechend sollte der Einsatz von KI-Methoden zielgerichtet, überlegt und immer im vollen Bewusstsein dieser Kosten erfolgen. Hierbei sollte unterschieden werden in tatsächliche Kosten (nicht nur im unmittelbaren monetären Sinne, auch Schäden an Umwelt, Mensch, Infrastrukturen), aktuell bestehende Risiken für Individuen und Institutionen (etwa durch Datenleaks, Systemüberlastung o. Ä.), und potenzielle Langzeitfolgen für die Informationslandschaft und Gesellschaft. Letztere lassen sich natürlich nie mit endgültiger Sicherheit vorhersagen; dennoch sollte man sie im Auge behalten.[2]

2. Kosten für Umwelt und Klima

LAI ist intrinsisch ressourcenintensiv, verbraucht große Mengen an Strom und Wasser und trägt damit massiv zur Beschleunigung des Klimawandels bei. So erforderte das Training des Modells GPT-4 von OpenAI ca. 50 Gigawattstunden (das entspricht der benötigten Energie für 200 Flüge von New York nach San Francisco) – 50mal mehr als für den Vorgänger GPT-3, und es ist fraglich, ob diese Werte die Experimentalphasen jeweils miteinschließen. Für das Training von GPT-3 wurden mindestens 5 Mio.

1 Dieser Beitrag basiert auf dem Vortrag mit gleichem Titel (Kasprzik, Argie: Die KI(rche) im Dorf lassen – Wege zu einer klima- und sozialverträglichen Nutzung von Künstlicher Intelligenz, Online: https://nbn-resolving.org/urn:nbn:de:0290-opus4-198413) am 26.06.2025 beim 9. Bibliothekskongress 2025 (zugleich 113. BiblioCon) in Bremen.), welcher wiederum eine Aufarbeitung eines digitalen Theaterstücks darstellt (Bach, Nicolas; Kasprzik, Argie: Einblick in die erste vollautomatische KIbliothek. Mit KI auf neuen Wegen in der Stadtbibliothek Bad Turing, 8:31 min, #vBIB24 – Digitale Teilhabe, Berufsverband Information Bibliothek e. V. (BIB) et al., 04.12.2024. https://doi.org/10.5446/69373). Die Vortragsfolien und der vorliegende Beitrag verweisen auf eine Reihe von Onlinequellen – aufgrund der Dynamik der Entwicklungen und der Tatsache, dass einige Links subjektive Positionen aus diversen Web-Communities wiedergeben, möchte ich die Lesenden jedoch nachdrücklich dazu ermuntern, ergänzend eigene Recherchen durchzuführen.

2 Interessierten Lesenden seien zur Vertiefung u.a. die Begriffe der „Technikfolgenabschätzung" und der „Sozioinformatik" samt zugehöriger Literatur empfohlen.

Liter Wasser aufgewendet und für das des Modells LLaMa-3 von Meta 22 Mio. Liter – das entspricht z. B. der benötigten Menge für den Anbau von 2 bzw. 4,5 Tonnen Reis, die dann u. a. für Nahrungserzeugung und Dürrebekämpfung fehlt. Während die Trainingsphase eines Modells zeitlich begrenzt ist, erfordert der Betrieb darauf aufsetzender Anwendungen fortlaufend weitere Ressourcen, und diese summieren sich – manche LAI-Datenzentren verbrauchen ähnlich viel Strom und Wasser wie eine mittelgroße Stadt. Es gibt Prognosen, dass LAI bis 2030 ca. ein Fünftel des globalen Netzbetriebes und jährlich eine Strom- und Wassermenge in Anspruch nehmen wird, die den Verbrauch eines kleineren Staates wie z. B. Wyoming oder Dänemark um ein Fünffaches übersteigt. Die Energieerzeugung für LAI-Datenzentren aus fossilen Quellen vermehrt die Luftverschmutzung, was wiederum Atemwegserkrankungs- und Todeszahlen steigert (Prognose 2030 für die USA: durch LAI viermal mehr Todesfälle als noch 2023) und wovon infrastrukturbedingt besonders ärmere Communities betroffen sind. Durch Training und Betrieb erhöht sich der Bedarf an Hardware und damit fällt auch mehr Sondermüll an: Je nach Szenario könnte LAI bis 2030 pro Jahr bis zu 2,5 Mio. Tonnen Elektronikschrott produzieren – das entspricht etwa einem weggeworfenen Smartphone pro Mensch. Wenn man die Betriebskosten (stark vereinfacht) herunterbricht auf einzelne Interaktionen mit Anwendungen wie ChatGPT, dann kostet eine LAI-Chatbot-Anfrage ca. 3 Wh und damit sechs- bis zehnmal mehr Strom als eine prä-LAI-Google-Suche, das Generieren einer 100-Wort-Email etwa soviel Strom wie 7 iPhone-Aufladungen und 0,5 Liter Wasser, eine Chatbot-Unterhaltung mit 10 bis 50 Anfragen grob 2 Liter Wasser.[3]

Bei all diesen Angaben ist es wichtig darauf hinzuweisen, dass dies natürlich nur Schätzungen sein können. Sie sollen eine Vorstellung der Dimensionen vermitteln, keine falsche Präzision vortäuschen. Da sich die Rahmenbedingungen fortlaufend ändern, ändern sich auch die Prognosen – allerdings, obwohl sich einige dieser Größen mit vorhandenem Willen in den Konzernen und in der Politik sicher reduzieren ließen, scheinen sie im kommerziellen Bereich mehrheitlich noch ungebremst zu steigen.

3. Ökonomische und psychische Kosten

Es ist weiterhin so, dass viele LAI-Ansätze angewiesen sind auf von Menschen annotierte bzw. korrigierte Daten, und entsprechend wenig darf diese menschliche Arbeit aus der Sicht großer Konzerne kosten. So wird etwa die Kontrolle von Inhalten auf Social-Media-Plattformen zu großen Teilen über Subunternehmen aus den Industriestaaten abgeschoben in weit weniger regulierte Regionen wie etwa Kenia oder Argentinien, für einen Hungerlohn und unter krankmachenden Bedingungen. Sogenannte Contentworker*innen sortieren täglich viele Stunden beispielsweise Beiträge mit Darstellungen von Gewalt und Tod, entscheiden über deren Löschung und produzieren so wertvolle Trainingsdaten für KI-basierte Modelle zur inhaltlichen Qualitätssicherung.[4] Eine längerfristige Ausübung dieser Tätigkeit führt häufig zu schwerwiegenden gesundheitlichen und psychischen Schäden – gegen Facebooks Mut-

3 Es gibt viele Versuche, diese Kosten zu quantifizieren. Die obigen Aussagen beruhen auf den folgenden Quellen (alle Stand: 01.09.2025): AI needs your help, https://savethe.ai/; Harper, Christopher: Using GPT-4 to generate 100 words consumes up to 3 bottles of water, 19.09.2024, https://www.tomshardware.com/tech-industry/artificial-intelligence/using-gpt-4-to-generate-100-words-consumes-up-to-3-bottles-of-water-ai-data-centers-also-raise-power-and-water-bills-for-nearby-residents; Edwards, Benj: AI in Wyoming may soon use more electricity than state's human residents, 29.07.2025, https://arstechnica.com/information-technology/2025/07/ai-in-wyoming-may-soon-use-more-electricity-than-states-human-residents/.

4 Siehe z. B. Dachwitz, Ingo: Die versteckten Arbeitskräfte hinter der KI erzählen ihre Geschichten, 08.07.2024, https://netzpolitik.org/2024/data-workers-inquiry-die-versteckten-arbeitskraefte-hinter-der-ki-erzaehlen-ihre-geschichten/, Stand: 01.09.2025.

terkonzern Meta und das Subunternehmen Sama läuft seit 2023 eine Massenklage von über 180 kenianischen Contentworker*innen aufgrund massiver posttraumatischer Störungen und menschenunwürdiger Arbeitsbedingungen. Beide Unternehmen wiesen die Vorwürfe zurück.[5]

4. Weitere gesellschaftliche und individuelle Kosten und Risiken

Ein Einstieg in LAI ist aufgrund der erforderlichen Investitionen mit beträchtlichen Hürden versehen. Entsprechend ist ein kleiner Kreis von dominierenden Techfirmen in der Lage, in hohem Tempo und bisher ohne wesentliche Regulierung ein Oligopol aufzubauen und so zu „Gatekeepers" zu werden. Sie haben es beispielsweise über die Preisgestaltung in der Hand, gewisse Funktionalitäten ihrer LAI-Anwendungen nur solchen Individuen und Institutionen zugänglich zu machen, die sich diese (in Form von Geld oder Daten) leisten können. Über Kooperationsvereinbarungen können sie ggf. auch kontrollieren, ob der intendierte Verwendungszweck mit ihren wirtschaftlichen und politischen Prioritäten übereinstimmt. Das räumt ihnen Möglichkeiten zur Einflussnahme in einer Reihe von Einsatzszenarien ein, die im öffentlichen Interesse liegen und daher neutral gehandhabt werden sollten, etwa in den Bereichen Forschung, Bildung und Infrastruktur.[6] Ein weiteres, ganz konkretes Problem sind Serviceunterbrechungen durch sogenannte AI Crawler, die im großen Stil und teilweise trotz technischer Gegenmaßnahmen Daten abziehen und damit Webserver lahmlegen – das schließt auch Open Repositories und andere am Gemeinwohl orientierte Informationsinfrastrukturen mit ein.[7]

Nicht nur für Webseitenbetreibende, auch für deren Nutzende ist es aktuell häufig nicht ersichtlich, für welche LAI-Zwecke ihre Daten genutzt werden und an welchen Stellen sie mit genKI-Anwendungen interagieren. Beispielsweise verwendete OpenAI Beiträge aus einem Diskussionsforum, um Chatmodelle mit besseren Überzeugungsfähigkeiten zu trainieren, und testete deren Output in einer kontrollierten Umgebung. Aufgrund der öffentlichen Verfügbarkeit der Daten war das legal, rief aber bei einigen Usern Empörung hervor, da sie über diese Art der Nutzung gerne vorab informiert worden wären. Eine Gruppe von Forschenden der Universität Zürich führte ähnliche Experimente durch, speiste die Resultate jedoch auch wieder als Diskussionsbeiträge in das Forum ein, und zwar entgegen den Forumsregeln ohne sie als KI-generiert zu markieren und ohne Zustimmung der Betreibenden. Die Forschenden bekamen in diesem Fall eine Verwarnung und durften die Studie nicht publizieren. Dennoch illustrieren Vorkommnisse dieser Art, wie genKI die Herausforderung noch vergrößert, Transparenz in Bezug auf die Herkunft und Verwendung von Daten im Netz herzustellen – ganz zu schweigen davon, dass sie zumindest aus der Sicht einiger Nutzender solcher Gesprächsforen das Potenzial hat, die aufrichtige Diskussionskultur endgültig zu zerstören („nach Troll kommt Bot").[8]

5 Booth, Robert: More than 140 Kenya Facebook moderators diagnosed with severe PTSD, 18.12.2024, https://www.theguardian.com/media/2024/dec/18/kenya-facebook-moderators-sue-after-diagnoses-of-severe-ptsd, Stand: 01.09.2025.

6 Siehe u. a. Abschnitt „Firmenkunden als wichtigste Zielgruppe" von Book, Simon; Hoppenstedt, Max: OpenAI stellt neue ChatGPT-Version vor, DER SPIEGEL (online), 07.08.2025, https://www.spiegel.de/netzwelt/chat-gpt-5-openai-veroeffentlicht-neue-version-a-80fd70da-fba9-46bf-bb93-0bac64d58dce und Henning, Maximilian: Monopole verhindern und das Gemeinwohl fördern, 20.10.2024, https://netzpolitik.org/2024/kuenstliche-intelligenz-monopole-verhindern-und-das-gemeinwohl-foerdern/, beide Stand: 01.09.2025.

7 Shearer, Kathleen; Walk, Paul: The impact of AI bots and crawlers on open repositories. Results of a COAR survey, April 2025, Confederation of Open Access Repositories (COAR), 03.06.2025. Online: https://coar-repositories.org/news-updates/open-repositories-are-being-profoundly-impacted-by-ai-bots-and-other-crawlers-results-of-a-coar-survey/, Stand: 01.09.2025.

5. Risiken für Informationsqualität und menschliche Kompetenzen

Risiken ergeben sich auch dann, wenn LAI-Output oder LAI-Dienste vorschnell und ohne die nötigen Hintergrundinformationen verbreitet werden, bevor sie einen ausreichenden Grad der Reife erreicht haben. Das Netz füllt sich mit sogenanntem „AI slop", also KI-generierten Text- und Bildbeiträgen niedriger Qualität. Dadurch wird es noch schwieriger, Inhalte mit Substanz zu finden (und solche ohne Substanz zu erkennen) – und es hat möglicherweise sogar Folgen für zukünftige auf Webinhalten trainierte genKI-Modelle, denn synthetische Trainingsdaten können sich negativ auf die Performanz von Modellen auswirken.[9]

Erschwerend kommt hinzu, dass viele Nutzende weiterhin nicht ausreichend über die Funktionsweise von LAI-Anwendungen wie ChatGPT und Co. informiert sind und dadurch mit falschen Erwartungen an diese herangehen. So deuten vermenschlichende Ausdrucksweisen wie „die KI halluziniert" auf die intuitive Annahme hin, dass die Anwendung (zusätzlich zu einer menschenähnlichen Mitteilungsabsicht) einen unmittelbaren Zugang zu gängigem Weltwissen hat.[10] Die durch statistische Häufungen in den Trainingsdaten erzielten „Glückstreffer" müssen jedoch mit Hilfe verschiedener Ansätze ergänzt werden, um so das Wissen mühsam zu simulieren, etwa durch die Einbindung von explizit dokumentierten Informationen (Retrieval Augmented Generation; RAG) und durch Schlussmechanismen (Reasoning). Da das bisher nur sehr unvollständig gelingt, empfiehlt es sich in Szenarien wie z. B. der Softwareentwicklung, in denen Korrektheit und Präzision essenziell sind, die Ergebnisse nicht ungeprüft weiterzuverwenden – und je komplexer der Gegenstand der Anfrage, desto mehr Fachkenntnis ist für diese Prüfung erforderlich. Gezielt eingesetzt und mit menschlicher Expertise verzahnt hat genKI durchaus das Potenzial, eine ganze Reihe von Arbeitsprozessen zu transformieren. Wenn bei Nutzenden und auch bei Arbeitgebenden jedoch die Annahme besteht, dass Fachkompetenzen durch genKI ersatzlos ausgetauscht werden können (und dem von den Anbietenden dieser Lösungen bewusst nicht widersprochen bzw. Übertreibung als Verkaufsargument genutzt wird), steigt das Risiko negativer Folgen signifikant – von unbemerkten Fehlern im Prozess und entsprechend schlechtem Output[11] bis hin zu den destruktiven Auswirkungen eines verfehlten Change Managements auf Psyche und Produktivität der Mitarbeitenden.

8 Ferguson, Mackenzie: OpenAI Turns r/ChangeMyView into AI's Persuasion Playground! 01.02.2025, https://opentools.ai/news/openai-turns-rchangemyview-into-ais-persuasion-playground; Travis, Kate: AI-Reddit study leader gets warning as ethics committee moves to 'stricter review process', 25.04.2025, https://retractionwatch.com/2025/04/29/ethics-committee-ai-llm-reddit-changemyview-university-zurich/, beide Stand: 01.09.2025.

9 Alemohammad, Sina; Casco-Rodriguez, Josue; Luzi, Lorenzo u. a.: Self-Consuming Generative Models Go MAD, 2023. https://doi.org/10.48550/arXiv.2307.01850.

10 Und das, obwohl sich eine künstliche allgemeine Intelligenz auf dem Niveau eines menschlichen Intellekts bisher nicht manifestiert hat – Prognosen schwanken von „2027" (Kokotajlo, Daniel; Alexander, Scott; Larsen, Thomas u. a.: AI 2027, 03.04.2025, https://ai-2027.com/, Stand: 01.09.2025) bis „in weiter Ferne".

11 Das gilt ebenso für schlechte menschliche Arbeit, aber mit genKI lassen sich nun mit wenig Aufwand viel größere Mengen an professionell wirkendem Material generieren, für das sowohl die Schwelle eines Anfangsverdachts und damit der Auslöser als auch der Aufwand für eine Prüfung entsprechend höher liegen. Siehe u. a. (beide Stand: 01.09.2025) Aldridge, David: The case against vibe coding, 12.06.2025, https://www.theserverside.com/tip/The-case-against-vibe-coding und Diskussion unter Dangerous_Ad_2357, Vibe Coding is killing my company, 10.07.2025, r/vibecoding, Reddit, https://www.reddit.com/r/vibecoding/comments/1lwbgzl/vibe_coding_is_killing_my_company/.

Auch in der Bildung und Forschung, einem für wissenschaftliche Bibliotheken zentralen Bereich, stellt sich die Frage nach den Auswirkungen einer extensiven Nutzung von LAI-Anwendungen auf den wissenschaftlichen Prozess – insbesondere, wenn dieser gerade erst gelernt wird. Erste, allerdings nichtrepräsentative Studien deuten auf das Risiko eines Verfalls der eigenen Expertise bzw. Fähigkeit zum selbstständigen Denken („kognitive Schulden") hin, wenn Menschen sich bei gewissen Aufgaben (z. B. Essay-Schreiben, Krebserkennung) ausschließlich auf LAI-Unterstützung verlassen.[12] Ob diesen Schulden entgegengewirkt werden kann, indem beim Entwickeln und Erhalten intellektueller Fähigkeiten konsequent darauf geachtet wird, LAI-Werkzeuge nur mit Augenmaß und gezielt einzusetzen und sich dabei einen kritischen Blick zu bewahren, wird sich wohl erst über die nächsten Jahrzehnte herausstellen. Sicher ist jedoch, dass die Suche nach geeigneten Herangehensweisen an diese Werkzeuge Menschen in Bildung und Forschung noch über viele Jahre beschäftigen wird.

Wie anfangs angekündigt, werden in den obenstehenden Abschnitten neben bereits belegbaren Kosten auch eine Reihe potenzieller Langzeitschäden genannt, die nicht zwingend genau so eintreten müssen oder sich zumindest abmildern lassen. Eine mündige Gesellschaft entlastet das jedoch nicht von der Aufgabe, diese Kosten und Risiken im Detail zu kennen und sich auf institutioneller, nationaler und internationaler Ebene dafür zu engagieren, dass diesen aktiv und langfristig entgegengewirkt wird. Das gilt in ganz besonderem Maße für die Bibliothekscommunity – schließlich liegt es neben der Pflicht zu ethischem und nachhaltigem (und somit ökologischem und sozialbewusstem) Handeln im Kernauftrag von Bibliotheken, das Bestehen frei zugänglicher und sinnvoll nutzbarer Information zu sichern und voranzutreiben.

6. Der Gegenentwurf

Osma Suominen, Spezialist für Informationssysteme und Schöpfer des Open-Source-Machine-Learning-Toolkits Annif für automatisierte Inhaltserschließung, stellte bei einer Veranstaltung namens „AI Sauna" der Finnischen Nationalbibliothek in einem Impulsvortrag die folgenden fünf Prinzipien[13] dafür auf, wie man sich dem Thema KI ethisch und nachhaltig nähern kann:

1. use AI to make the world better
Dies geht offensichtlich gut zusammen mit dem öffentlichen Auftrag von Bibliotheken.

2. use the smallest AI that works
Es muss gar nicht immer LAI sein – vielleicht genügt bereits eine unaufwändigere Form der Automatisierung wie etwa ein Skript, oder eben ein kleines statistisches Machine-Learning-Modell oder für

12 Diskussion zu potenziellen Kompetenzverlusten in der Bildung siehe Reinmann, Gabi: Deskilling durch Künstliche Intelligenz?, in: Hochschulforum Digitalisierung, Diskussionspapier Nr. 25, Oktober 2023. Online: https://hochschulforumdigitalisierung.de/sites/default/files/dateien/HFD_DP_25_Deskilling.pdf, Stand: 01.09.2025. Studien: Kosmyna, Nataliya; Hauptmann, Eugene; Yuan, Ye Tong u. a.: Your Brain on ChatGPT. Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task, arXiv, 2025. https://doi.org/10.48550/arXiv.2506.08872 und Budzýn, Krzystof; Romańczyk, Martin; Kitala, Diana u. a.: Endoscopist deskilling risk after exposure to artificial intelligence in colonoscopy, in: Lancet Gastroenterol Hepatol, 12.08.2025. https://doi.org/10.1016/S2468-1253(25)00133-5.

13 Suominen, Osma: Building Civilized AI, Minute 22:00 bis 31:00 von AI Sauna, 06.05.2024. https://www.youtube.com/watch?v=oT8FP1JH5vE&t=1322s. Folien online: https://docs.google.com/presentation/d/e/2PACX-1vRA1o11pODoJ0FmFc8dRj-xNZRUs7lsxzDACkiYt6d-Bdfql1ujw3gGpSedTQnXDG0MrRg3_WAI1GQS/pub, beides Stand: 01.09.2025.

sprachbezogene Anwendungsfälle ein Small Language Model. Letztere sind leichtgewichtiger, können lokal trainiert und betrieben werden, und sind für spezifische Aufgaben oft ebenso gut geeignet wie Large Language Models.[14]

3. don't depend on corporate AI

Techfirmen handeln primär im eigenen und nicht im öffentlichen Interesse, mit allen Nachteilen, die sich daraus ergeben. Auf kommerziellen Lösungen aufgebaute Systeme sind von deren Anbietenden und deren Modellauswahl abhängig – es können nicht wie im Open-Source-Bereich einfach andere Komponenten gewählt werden, wenn diese Lösungen nicht mehr den gewünschten (inhaltlichen, finanziellen, ethischen) Bedingungen entsprechen.

4. evaluate & create data sets

Hier ergibt sich ein perfektes Tätigkeitsfeld für Bibliotheken. Zunächst einmal werden KI-Modelle besser, wenn sie mit hochqualitativen, unverzerrten Daten trainiert sind. Darüber hinaus brauchen wir aber auch möglichst diverse und domänenspezifische Datensammlungen, um Modelle auf unsere eigenen Zwecke hin zu testen, zum Beispiel hinsichtlich ihrer Performanz in verschiedenen Sprachen und Kulturen, und um dann ggf. auf der Basis dieser Daten ein geeignetes Modell für einen konkreten Anwendungsfall durch weiteres Training „nachzujustieren" (Finetuning).

5. be open and transparent

Sich für mehr Openness – Open Access, Open Data, Open Science und zunehmend auch Open Source – einzusetzen ist naturgemäß eine zentrale Aufgabe für Bibliotheken. Kommerzielle Unternehmen im Techbereich haben aus Wettbewerbsgründen wenig Interesse daran, ihre Methoden offenzulegen. Um aber LAI-basierte Systeme trotz des intrinsischen Blackbox-Charakters eines Machine-Learning-Ansatzes auch nur annäherungsweise nachvollziehbar und nachnutzbar zu machen, müssen mindestens die folgenden Elemente bekannt gemacht werden: die für das Training verwendeten Daten, die Modellarchitektur, die beim Training gesetzten Parameter (insbesondere die Gewichtungen) und der Quellcode zum Training und Betrieb des Systems. Nur so können solche Lösungen reproduziert und an andere Zwecke angepasst werden. Einige Techfirmen wie z. B. Meta bewerben ihre Modelle als „Open Source", ernten dafür jedoch heftige Kritik aus der Open-Source-Community, da die obenstehenden Bedingungen eben nicht alle erfüllt sind. Meta und OpenAI legen bei ihren neuesten Modellen LLaMa-4 und GPT-OSS nun zwar die Trainingsgewichte offen, beschränken aber weiterhin den Zugriff auf Trainingsdaten und Quellcode. Im Gegensatz dazu fallen Projekte von öffentlichen Institutionen wie etwa ein von den beiden Eidgenössischen Technischen Hochschulen Zürich und Lausanne angestoßener Zusammenschluss von ca. 10 Schweizer akademischen Einrichtungen und über 800 Forschenden zu „SwissAI" positiv auf. Dieses Projekt soll im zweiten Halbjahr 2025 in der Veröffentlichung eines „Large Language Model für das Gemeinwohl" einschließlich der Trainingsdaten, des Quellcodes und der Modellgewichte münden.[15]

14 Nragi, David: Small Language Models vs. Large Language Models. Understanding the Differences and Implementations, 30.07.2024, https://medium.com/@nagidravid/small-language-models-vs-large-language-models-understanding-the-differences-and-implementations-fc91ff208541, Stand: 01.09.2025.

7. Als Institution aktiv werden

Natürlich kann jede Institution nur im Rahmen ihrer Möglichkeiten agieren, sowohl intern als auch nach außen. Intern bietet sich die Erarbeitung und Umsetzung einer KI-Strategie (oder generell Nutzungsstrategie für Soft- und Hardware) an, die die oben genannten Aspekte berücksichtigt und ressourcenschonendere Lösungen priorisiert. Nach außen können Bibliotheken ihrer Multiplikatorenrolle gerecht werden, indem sie ihre Nutzenden und die Allgemeinheit über die Kosten und Risiken einer extensiven Nutzung von LAI aufklären. Auf der politisch-strategischen Ebene ist die logische Konsequenz aus dem Kernauftrag und den Leitsätzen von (insbesondere wissenschaftlichen) Bibliotheken übertragen auf LAI ein fortgesetzter Aktivismus zur Förderung von Openness und zur Verbesserung der Rechtslage in Richtung *fair use*. Das bedeutet, es muss gesetzlich sichergestellt werden, dass öffentlich geförderte Einrichtungen für Zwecke der gemeinwohlorientierten Forschung und Informationsversorgung freien Zugriff auch auf urheberrechtsgeschützte Datensammlungen zum Training von Machine-Learning-Modellen haben, während profitgesteuerten Akteuren wie etwa Techfirmen und Verlagen dieses Recht nicht eingeräumt werden sollte. Um Projekte zur Trainingsdatenakquise und zum Entwickeln und Betreiben von frei und offen nutzbaren, ressourcenschonenden, datenschutzbewussten und ethisch ausgerichteten KI-Anwendungen dann auch erfolgreich umzusetzen, führt kaum ein Weg am Aufbau von Kooperationen zwischen möglichst vielen Einrichtungen vorbei. Dabei kommt großen Kompetenzzentren eher eine konzeptionelle und koordinierende Rolle zu, während kleine, nachnutzende Bibliotheken beispielsweise über Datenlieferungen und Praxistests zuarbeiten können.

8. Als Individuum bewusst handeln

Die oberste Maxime lautet: Wer zu einer mündigen Entscheidung kommen können will, muss stets selbst prüfen. Das gilt für die Ausführungen in diesem Beitrag (recherchieren Sie gerne nochmal den neuesten Stand!), es gilt für jeglichen KI-generierten Output in allen Verwendungskontexten, in denen Faktentreue eine Rolle spielt, und es gilt z. B. auch für die Rechte, die Sie den Anbietenden einräumen müssten, wenn wieder eine App drängelt, doch bitte diese und jene KI-gestützte Funktionalität zu aktivieren. Meine Empfehlung an Menschen, die sich angesichts der Allgegenwärtigkeit von LAI bewusster durch die digitale Landschaft bewegen wollen: Beobachten Sie Ihre eigene Praxis und passen Sie sie entsprechend an. Nicht jede Spielerei mit einem LAI-Dienst (etwa Bildgenerierung zu dekorativen Zwecken) ist notwendig und längst nicht jede Aufgabe erfordert den Einsatz eines LAI- oder auch nur KI-basierten Werkzeugs. D. h. statt sich für jede noch so triviale Fragestellung auf eine allumfassende LAI-Assistenz wie ChatGPT, Copilot & Co. zu verlassen und damit ein beträchtliches Stück Kontrolle abzugeben, wählen Sie Ihr Arbeitsmittel gezielt von Fall zu Fall. Häufig ist ein spezifisches Instrument nicht nur ressourcenschonender, sondern auch besser geeignet – der Taschenrechner, das Nachschlagen in einer Fachdatenbank, ein simples Skript oder Plugin, ein Spezialwerkzeug auf der Basis eines

15 Open Source Initiative: Open Source AI Definition version 1.0., Oktober 2024, https://opensource.org/ai/open-source-ai-definition. Zu Metas Open Washing siehe Paul: Openness is binary, forkable, 11.04.2025, https://www.forkable.io/p/metas-new-llama-4-ai-models-arent#§openness-is-binary. Zu SwissAI siehe Meyer, Florian; Anchisi, Mélissa: A language model built for the public good, ETHZ, 09.07.2025, https://ethz.ch/en/news-and-events/eth-news/news/2025/07/a-language-model-built-for-the-public-good.html, alle Stand: 01.09.2025.

an die Aufgabe angepassten Machine-Learning-Modells, … Zugegebenermaßen erfordert das im Alltag aufgrund der andauernden Bedrängung durch Anbietende, Medien und Peergruppen, doch lieber die neueste kommerzielle Trendlösung zu nutzen, zusätzlichen mentalen Aufwand, und ganz konkret den Aufbau einer gewissen Toolkompetenz. Fangen Sie im Zweifel klein an – wenn Sie z. B. das Auslösen der LAI-basierten Zusammenfassung bei Google und den damit verbundenen Strom- und Wasserverbrauch vermeiden wollen, steuern Sie die klassische Websuche von Google direkt an.[16] Suchen Sie in allen von Ihnen genutzten digitalen Anwendungen in den Einstellungen nach LAI-Funktionalitäten und wählen Sie diese wo immer möglich ab, wenn Sie sie nicht benötigen. Achten Sie auf Ihre eigenen Rechte – wenn etwa jemand in einem Videocall vorschlägt, die KI-Assistenz zu aktivieren, um eine Zusammenfassung des Besprochenen zu erhalten, Sie aber nicht wollen, dass Ihre Stimme, Ihr Videobild und Ihre Chatnachrichten mitverarbeitet werden, dann widersprechen Sie deutlich. Und tragen Sie die Botschaft weiter – auch an Ihre Leitungsebenen. Jedes gedankenlose Nutzen von LAI-Anwendungen signalisiert einer Techfirma, dass es sich lohnt, ihr Angebot weiter aufzublähen, obwohl der Bedarf ggf. gar nicht besteht.

9. Fazit

Der Aufruf dazu, bei der Nutzung von LAI genauer hinzuschauen, wird oft als „Maschinensturm“ diskreditiert und damit – wie der historische Maschinensturm auch – als Ablehnung der Automatisierung an sich oder auch dieser spezifischen Technologie missverstanden. Dass Transformerarchitekturen und darauf aufsetzende Ansätze wie Large Language Models und generative KI einen wissenschaftlichen Durchbruch darstellen, ist unbestritten. Wie alle wissenschaftlichen Durchbrüche sollten sie jedoch dazu genutzt werden, die Allgemeinheit voranzubringen, und nicht zur Bereicherung weniger. Und die Frage danach, was es auf der anderen Seite die Allgemeinheit kostet, muss stets ein zentrales Entscheidungskriterium für ihren Einsatz sein.

Der Blackbox-Charakter insbesondere der kommerziell angebotenen Lösungen führt dazu, dass für die breite Bevölkerung weder das ganze Ausmaß der Auswirkungen auf Klima und Gesellschaft noch die Funktionsweise der Anwendungen ausreichend transparent ist. Ersteres hat zur Folge, dass an den entscheidenden Stellen zu wenig gegengesteuert wird. Letzteres vergrößert das Risiko einer digitalen Abhängigkeit und Unmündigkeit, wenn Menschen nicht über ausreichend eigene fachliche Expertise für einen kontrollierten Einsatz der Technologie verfügen. Bibliotheken sind über dieses gesamte Spektrum hinweg in der Pflicht: aufzuklären, weiterhin Informationskompetenz zu fördern, auch bei der Suche nach Methoden zur Automatisierung der eigenen Aktivitäten ethisch und ökologisch zu handeln, und sich entschieden für einen systemischen Wandel hin zu mehr Openness einzusetzen.[17]

Argie Kasprzik, ZBW – Leibniz-Informationszentrum Wirtschaft, Hamburg, https://orcid.org/0000-0002-1019-3606

16 Siehe z. B. https://tenbluelinks.org/, Stand: 01.09.2025.

17 Der vorliegende Beitrag wurde vom Autor gänzlich ohne generative KI-Werkzeuge verfasst. Dieser Hinweis enthält keine Wertung bzgl. eines zweckmäßigen Einsatzes von KI-basierten Textwerkzeugen an sich und soll lediglich Transparenz in Bezug auf den Erstellungsprozess dieses konkreten Textes schaffen.

Zitierfähiger Link (DOI): https://doi.org/10.5282/o-bib/6201