

Sprechen erzeugt ein Glücksgefühl

Der Einsatz von KI im Bereich der Sprecher

geschrieben von Anna-Sophia Lumpe

»Die Illusion der Objektivität ist oft das Werkzeug derer, die die Macht haben, ihre eigenen Interessen zu verfolgen.« – Noam Chomsky.

So verhält es sich wohl, wenn Big Tech eine Lösung für das »Problem Sprecherinnen« aufzeigt. Dass fachfremde Personen Lösungen für Probleme anbieten, die gar nicht existieren, ist in der Neuzeit schon alltäglich. Dabei betrachten sie Prozesse aus ihrer Perspektive und bewerten sie nach ihren Parametern. Oft sind diese aber auf den vorliegenden Prozess nicht sinnvoll anwendbar, und es bedarf umfassenderer Grundlagen. Für ein zügiges Ergebnis werden diese aber übersprungen. So auch beim Einsatz von KI-Software in der Branche des Sprechens, wo der Arbeitsprozess selbst etwas ist, das unumkehrbar das Ergebnis beeinflusst, und der Moment der Aufnahme mit Menschen immer die Chance birgt, ein unerwartetes Ergebnis zu erzielen, das hochwertiger und ansprechender ist als eine digitale Lösung. Kreatives Schaffen ist nicht geradlinig, wenn es unmittelbar im Moment mit einer lebendigen Person stattfindet. Profis kommen auf Ideen, die nur durch Erfahrung und ganzheitliche Beherrschung des eigenen Instruments entstehen. Synthese bildet nur einen Mittelwert.

Doch KI-Unternehmen bewerten nach Zeit und Kosten. Essenzielles wird ignoriert. Stimme wird durch einen Klick mit starren Befehlen von Personen generiert, die ihre eigene nicht professionell einsetzen. Der notwendige Horizont zur Generierung fehlt: was möglich ist, was ausprobiert werden sollte, welche Emotion eventuell Gesprochenes überlagern kann, obwohl sie gegensätzlich zum Textinhalt steht, und vieles mehr. Das Ergebnis ist unkreativ. Aber auch dafür gibt es eine Lösung aus der Industrie: Die Sprecherinnen können die Generierung übernehmen! Sie dürfen ihren Ausdruck und ihre Interpretationen in Prompts übertragen und so die synthetische Stimmausgabe formen und ansprechend gestalten. Aber warum sollten wir statt mit unserem körpereigenen Instrument auf eine entmenslichte Art und Weise unsere Profession ausführen? Das Sprechen ist psychisch wie physisch erfüllend und erzeugt ein Glücksgefühl. Eine Intention mit dem eigenen Körper hörbar zu machen, führt zu einem Erfolgserlebnis. Es gibt Menschen, die über Denkprozesse wie korrekte Prompts für eine KI ein ähnliches Erfolgsgefühl erfahren, aber diejenigen, die mit ihren Emotionen arbeiten, empfinden so eben nicht. Folglich wäre es arrogant, nur aus einer fachfremden Perspektive und Welterfahrung heraus eine Profession und ihre Ausführung zu bewerten. Doch die aktuelle Situation führt zu Auftragsverlusten, da genau dieses Bewusstsein bisher fehlt.

Welche Chancen bietet KI für die sprechende Zunft? Bei der tatsächlichen Kreation wenig, denn wie bereits ausgeführt, wartet hier niemand auf den Tag, an dem die Ausführung automatisiert ist – bis auf ein paar Ausnahmen, bei denen die Stimme weniger belastet werden würde (z. B. Schreien, Krankheit). Viele Pflichten einer Selbstständigen haben allerdings wenig mit kreativer Arbeit zu tun, obgleich sie notwendig sind, um dem Beruf nachgehen zu können. Hier bietet KI viele Möglichkeiten, die verwaltenden Tätigkeiten, die Altersvorsorge, finanzielle Planung und die Akquise zu vereinfachen. Hier liegt das Potenzial, die Zeit für Kreativität zu vermehren und gleichzeitig den Schaffenden mehr Absicherung zu bieten. Eine automatische Verfolgung von lizenzierten Aufnahmen wäre eine große Entlastung und könnte gleichzeitig Vergütung gewährleisten, Urheberrechtsverletzungen vermeiden und das Persönlichkeitsrecht absichern. Größer als die Chancen sind allerdings derzeit die Bedrohungen, denn die Entwickler fragen aktuell nicht danach, wie KI die Kreativen unterstützen, sondern wie sie die Auftraggebenden (finanziell) entlasten kann. Dies resultiert in einer einseitigen, unvollständigen Aufklärung über die deutsche Rechtslage hinsichtlich Sprachaufnahmen und führt zu minderwertiger Qualität der synthetischen Ergebnisse. So werden auch zugrunde liegende Rechte wie Urheberrecht, Leistungsschutzrecht, Persönlichkeitsrecht und Datenschutz ignoriert und verletzt. Professionell Sprechende sind in Deutschland zwar rechtlich umfassend geschützt, aber staatliche Regelungen führen nicht dazu,

dass die Verletzung dieser Rechte wirklich unattraktiv wird, und die Betroffenen können sich nur mit Einsatz hoher finanzieller Mittel und viel Zeit gegen diese Rechtsbrüche wehren. Hier müsste unbedingt nachjustiert und garantiert werden, dass nicht nur große Verwertungskonglomerate mit hinreichend finanziellen Mitteln ihr Recht einfordern können, sondern vor allem die tatsächlichen Urheberinnen und Urheber: Einzelpersonen, deren Persönlichkeitsrecht verletzt wird.

Trainingsdaten von KI-Systemen müssen, wie im AI-Act angelegt, transparent gemacht werden, und bei Verstößen gegen Rechte kann nicht die Nachvergütung der Weg sein. Jede Person muss das Recht haben, der Nutzung ihrer biometrischen Daten (denn das ist die Stimme) uneingeschränkt widersprechen zu dürfen, sodass das KI-System diese entfernen und dann erneut trainiert werden muss. Nur so wird der Rechtsbruch unattraktiv.

»Der Einsatz von KI muss menschliche Entfaltung erweitern und darf sie nicht vermindern. KI darf den Menschen nicht ersetzen.« – Alena Buyx

Mehr zum Thema Künstliche Intelligenz in der Kultur in den Ausgaben [3/24](#), [4/24](#) und [5/24](#) von Politik & Kultur. Hier haben wir uns mit den Sparten Musik, Literatur und Übersetzung beschäftigt.

Dieser Text ist zuerst erschienen in [Politik & Kultur 6/2024](#).