

Bachelorarbeit

Über das Thema:

Sind KI-Stimmen ethisch vertretbar?

Timm Wehlte | 15474046

CC BY (4.0)

Erstprüfer: Prof. DIPL.-DOP Georgij Pestov

Zweitprüfer: DIPL.-KAM. Dominik Moos



Inhaltsverzeichnis

Einleitung	3
Wie KI funktioniert.....	4
Die Ethik des KI-Trainings	6
<i>Die Erstellung fremder Stimmen durch KI.....</i>	<i>6</i>
<i>Die Erstellung der eigenen Stimme durch KI</i>	<i>7</i>
<i>Die Erstellung neuer Stimmen durch KI</i>	<i>7</i>
Die Ethik der KI-Nutzung	9
<i>Die Nutzung fremder Stimmen durch KI</i>	<i>9</i>
<i>Die Nutzung der eigenen Stimme durch KI.....</i>	<i>15</i>
<i>Die Nutzung von neuen KI-Stimmen.....</i>	<i>19</i>
Conclusio	25
Literaturverzeichnis	27

Spätestens durch die plötzliche und für alle zugängliche Veröffentlichung von Chat-GPT, genauer gesagt GPT-3.5, im November 2022, ist Künstliche Intelligenz (KI) eines, wenn nicht sogar **das** große Thema unserer Zeit.

Für viele war Chat-GPT der erste Berührungspunkt mit einer gänzlich neuen Technologie, denn auch, wenn es bereits zuvor künstliche Intelligenz in verschiedenen Ausmaßen gab, von Sprachübersetzern bis hin zu "KI"-Figuren in Computerspielen, so war Chat-GPT noch einmal ein beeindruckender und faszinierender Sprung.

Ich weiß noch, wie ich bei meinen ersten Spielereien mit der Technologie überwältigt war von der Anzahl an Möglichkeiten, allesamt mit einzigartigen und authentischen Antworten. Zeilenlange Texte, Gedichte oder Aufsätze konnten in wenigen Sekunden entstehen, allesamt auf den ersten Blick stimmig, gut strukturiert und authentisch, was also steckt hinter dieser Technik?

Chat-GPT ist im Kern ein Large-Language-Model (LLM) und gilt damit auch als generative KI. Generative Künstliche Intelligenz beschreibt das Bundesamt für Sicherheit in der Informationstechnik wie folgt: *"[Generative KI-Modelle] werden mit großen Datenmengen trainiert und erlernen dabei Muster aus vorhandenen Daten. Daraus resultierend können sie neue Inhalte wie Texte, Bilder, Musik und Videos erzeugen, die ebenfalls diesen Mustern folgen"* (2025, BSI)

Auf die Veröffentlichung von Chat-GPT folgend gingen in kürzester Zeit etliche weitere Modelle online, wobei sich viele nicht nur auf Text beschränkten und anfangen Bilder, Videos, Musik oder Stimmen zu generieren. Anfangs noch sehr schwammig, aber in zeitlich kurzen Sprüngen immer realitätsnaher. Und mit diesen Modellen entstanden auch neue ethische Fragen, angefangen bei der Erstellung und dem "Trainieren" der Modelle bis hin zur letztendlichen Nutzung und Verbreitung KI-generierter Inhalte. Hinter verschlossenen Türen, vor allem im Silicon Valley, sind diese Fragen aber bei weitem nichts Neues und wurden schon vor Jahren diskutiert. Lange vor der Veröffentlichung von GPT-3.5 gab es innerhalb von OpenAI, der Firma hinter Chat-GPT und dem ehemaligen KI-Video Generator Sora, bereits beim Vorgängermodell GPT-2 intensive Debatten über die Gefahren der Technologie. *"If GPT-2 fell into the hands of terrorists, dictators, or clickbait farms they [OpenAI] reasoned, the model could be used for nefarious purposes... It was better to set a precedent for withholding research early. OpenAI, they decided, should not release the full version."* (Karen Hao, 2025, Empire of AI, S. 125)

Als (Chat-)GPT ungefähr 5 Jahre später dann doch in einer sogar noch besseren Version öffentlich zugänglich gemacht wurde, geschah dies nicht aus einer Re-Evaluierung der Risiken heraus, sondern aufgrund des Glaubens, dass die potenzielle, für OpenAI sogar unvermeidbare, Entwicklung einer Artificial General Intelligence (AGI) welche dem Menschen in quasi allen Bereichen ebenbürtig ist, deutlich größere Gefahren bietet: *"A language Model is a lot less powerful than an AGI. Im very worried about language models being weaponized for disinformation and this sort- that is very scary to me- but at the same time, it's a relatively singular and clear and defined concern"* (Zitat: Dario Amodei, ehemaliger Vice President of research bei OpenAI, Karen

Hao, 2025, Empire of AI, S.131)

Diese Aussage zeichnet ein gutes Bild der Ideologie von OpenAI und damit einhergehend auch einen Großteil der KI-Firmen in Silicon Valley: Jedes Risiko und jede potenzielle Gefahr ist nur ein unvermeidbarer und bedeutungsloser Fehltritt auf dem Weg zur ominösen "AGI".

Und auch wenn Chat-GPT vor allem Texte generiert, erfordert es keine große Abstraktion um diese Risiken und Gefahren, die auf Chat-GPT zutreffen, auch bei den Bild-, Video-, Musik- und Audio-Generatoren zu sehen. Eventuell sogar in einem noch größeren Ausmaß.

Diese Arbeit untersucht die Risiken und Chancen generativer KI um genau aufzuschlüsseln, was es mit diesen Thesen auf sich hat.

Hierfür wird erst beschrieben, wie ein KI-Modell funktioniert und trainiert wird, dann die Ethik des KI-Trainings analysiert, dann über die Ethik der Nutzung von KI, um dann schlussendlich abwägen zu können, wie ethisch vertretbar die Nutzung von Generativen-KI Modellen ist.

Ein spezifischer Fokus wird hierbei auf die Speech-Modelle gelegt, welche die Stimmen von bekannten oder erfundenen Menschen sowie fiktiven Charakteren imitieren und, mit selbst bestimmten Worten, wiedergeben können. Teile der Argumentation lassen sich aber auch sehr leicht auf andere Bereiche der Generativen-KI übertragen (Bild, Video, Text...) da die zugrunde liegende Technologie und wie sie funktioniert im Kern gleichbleibt.

Wie KI funktioniert

Die Möglichkeit Texte in Sprache umzuwandeln, wird von vielen Anbietern ermöglicht. ChatGPT, Gemini und Co. bieten mittlerweile die Möglichkeit direkt mit den Modellen zu reden, wodurch besagte Modelle eine menschennahe Stimme bekommen, mitsamt authentischem Sprachfluss, Redepausen oder sogar kleinen Lauten wie Lachen oder Atmen.

Dedizierte Anbieter, welche sich auf Text-to-Speech spezialisiert haben, liefern oft sogar noch bessere Ergebnisse und können dabei auch bereits existierende Stimmen "*klonen*": heißt sehr gut und teils schwer unterscheidbar imitieren. Beim Marktführer "ElevenLabs" geht das bereits für 6\$ im Monat. (ElevenLabs)

Um zu verstehen, wie diese Modelle funktionieren, wird diese Arbeit die Technologie dahinter auf das Fundamentale herunterbrechen. Die gesamte Technologie ist zwar deutlich komplexer und vielschichtiger, doch im Kontext dieser Arbeit reicht ein grobes Verständnis, um die ethischen sowie rechtlichen Implikationen zu verstehen.

Damit ein KI-Modell einen Output generieren kann, muss es erst einmal "trainiert" werden, wie uns der Medienanwalt und KI-Experte Sebastian Deubelli erklärt hat. Dafür werden Milliarden an Daten "gescraped", ein IT-Begriff, der das automatische Auslesen und potenzielle Extrahieren von Daten beschreibt. (Florian Ploner & Timm Wehlte, 2026, Bachelor Doku)

"Scraping" per se ist nicht illegal oder sogar unerwünscht, Suchmaschinen "scrapen" zum Beispiel Websites, um herauszufinden, ob diese dann eventuell relevante Informationen für den "Suchenden" haben. Diese Daten erhalten dann eine Bezeichnung, damit das Modell einordnen kann, um was genau es sich handelt. (Manfred Bremmer, 2026, Computerwoche)

So könnten Bilder eingeteilt werden in Tiere (Katze, Hund, Tiger...), während für die Modelle selbstfahrender Autos eher wichtig ist, Daten nach Verkehrsschildern oder anderen Fahrzeugen im Verkehr einzuteilen. An diesen Daten wird das Modell dann "trainiert", indem es Muster erkennt und dadurch sehr gut darin wird, diese wiederzuerkennen oder sogar wiederzugeben. (Abirami Vina, 2025, ultralytics)

Schaut man sich zum Beispiel Millionen Sätze aus Blogeinträgen, Romanen, Podcasts und so weiter an, ist es sehr wahrscheinlich, dass nach einem "Hallo wie geht es...", das Wort "dir" folgt, während auf den ganzen Satz "Hallo, wie geht es dir?" dann wiederum wahrscheinlich ein "Mir geht es gut" folgt.

Mit diesen Wahrscheinlichkeiten arbeiten KI-Modelle und können so erstaunlich echt wirkende Texte, aber eben auch Bilder, Videos und Stimmen erstellen.

Laut NVIDIA, einer der zurzeit wertvollsten und wichtigsten Firmen im KI-Bereich, sind für die Daten vor allem 2 Aspekte wichtig: Qualität und Diversität. (NVIDIA, What Is Generative AI?)

Qualität bedeutet, dass Bild-Daten klar erkennbar und in hoher Auflösung sind, Text-Daten gut geschriebene und echte Texte sind und Stimmen-Daten im Besten Fall in einem Studio mit einem High-End Mikrofon aufgenommen werden. Letztendlich kann ein KI-Modell nur die Qualität seiner Trainingsdaten erreichen, sind alle Stimmen-Daten windig und voller Rauschen, so wird das Modell auch am Wind und dem Rauschen trainieren und diese Verzerrungen zwangsläufig auch wieder ausspucken. Wichtig ist hier, dass eine KI eben nichts "versteht", wie ein Mensch oder sogar ein Tier es tut. Ein für uns klar erkennbares Störgeräusch ist für ein KI-Modell im Grunde dasselbe wie eine wunderschöne Stimme, umso wichtiger sind hochwertige Daten.

Diversität bedeutet, dass der Datensatz eine Vielzahl an Variationen bietet, um ein einseitiges Bias zu verhindern. Um ein Bild einer Katze zu generieren, braucht es nicht nur einfach hochauflösende Bilder von Katzen, es braucht diese Bilder von allen Seiten, von schwarzen sowie weißen Katzen, von großen und kleinen Katzen, von Katzen mit Behinderungen und so weiter. Bei einer Stimme braucht es die Stimmen aller Geschlechter, Emotionen, Akzente, verschiedene Tiefen und Höhen. Um die gigantische Vielzahl der Welt zu imitieren, braucht es auch eine gigantische Vielzahl an Daten, die diese widerspiegeln.

Exakt dieses Training ist der erste wichtige Faktor, welcher für die Leitfrage dieser Arbeit betrachtet wird

Die Ethik des KI-Trainings

Das Bewerten der Trainingsdaten ist sehr schwierig, denn es gibt verschiedene Möglichkeiten eine KI zu trainieren. Je nachdem welches Ziel man verfolgt, braucht man auch verschiedene Mengen und Arten von Daten, weshalb das Training im Folgenden in 3 Kategorien eingeteilt wird, um diese dann separat zu besprechen und zu beurteilen.

Die Erstellung fremder Stimmen durch KI

Als Erstes geht es um das Erstellen eines KI-Klons einer fremden Stimme. Dabei kann es sich um einen Prominenten handeln, vom Politiker bis zum Schauspieler, aber auch um private Kontakte. Plattformen wie ElevenLabs bieten an, Stimmen zu klonen und dann frei zu verwenden. Hierfür wird die KI zuerst mit Daten gefüttert, wobei wieder die 2 Stützpfeiler Qualität und Diversität gelten, heißt: desto hochwertiger und diverser die Daten sind, umso überzeugender ist das letztendliche Ergebnis.

Laut ElevenLabs reichen für einen simplen Klon mittlerweile schon 1-5 Minuten an Audio, für einen professionellen Klon braucht es über 30 Minuten. (ElevenLabs)

Die Anwendungsmöglichkeiten hierfür sind mehr oder weniger unendlich. Von harmlosen Parodien bis hin zu gefährlichen Fake News ist alles möglich. Auch lassen sich so Texte einer bekannten Persönlichkeit "authentisch" vorlesen, fiktive Figuren können auch nach dem Tod der sprechenden Person ihre Stimme behalten und so weiter.

Über die Ethik dahinter geht es im Kapitel **Die Ethik der KI-Nutzung**.

Das Klonen einer fremden Stimme ist ein heikles Thema, da man es theoretisch jederzeit machen könnte, auch wenn die betroffene Person überhaupt nichts davon weiß. Gleichzeitig ist es aber auch möglich, alle wichtigen Rechte mitsamt dem Einverständnis der Person einzuholen. ElevenLabs hat zum Beispiel Deals mit Schauspielern wie Michael Caine abgeschlossen, um nutzbare KI-Klone ihrer Stimme zu kreieren. Somit ist nicht nur rechtlich die Nutzung geklärt, auch profitieren die Schauspieler von den KI-Klonen durch eine neue Einnahmequelle. Caine sagte dazu: *"ElevenLabs is at the very forefront of technology, using innovation not to replace humanity, but to celebrate it"* (Dustin Blank, 2025, ElevenLabs)

Auch wurden Verträge mit den Estates weiterer bekannter, jedoch bereits verstorbener Persönlichkeiten abgeschlossen, um so zum Beispiel die KI-Klone von Judy Garland, Robert Oppenheimer oder John Wayne zu ermöglichen. Laut ElevenLabs würden diese Projekte zeigen: *"how responsible, high-quality AI voices can preserve the emotion of classic works while making them more accessible and engaging for listeners everywhere"* (Dustin Blank, 2025, Elevenlabs)

Das Erstellen fremder Stimmen durch KI ist also grundsätzlich nicht als ethisch oder unethisch zu bezeichnen, erst bei der Nutzung lässt sich hierbei ein begründetes Urteil fällen.

Die Erstellung der eigenen Stimme durch KI

Als nächstes geht es um die Erstellung der eigenen Stimme mithilfe eines KI-Klons. Eine Anwendungsmöglichkeit hierbei wäre, dass man Inhalte gar nicht mehr Einsprechen muss, was Zeit und Geld sparen kann. Ein Text lässt sich in wenigen Minuten von der KI vorlesen und somit fallen auch die Kosten für teures Equipment weg. Theoretisch braucht man nur einen Tag in einem Studio und könnte danach komplett auf ein Mikrofon verzichten. Somit ist man auch sehr flexibel, wenn es ums Reisen geht, denn egal wo man ist, man kann die eigene Stimme jederzeit generieren.

Da die Daten für das Erstellen eines KI-Klons von der eigenen Stimme auch die eigenen Daten sind, macht es das ganze ethisch sehr eindeutig. Was man mit seiner eigenen Stimme macht und ob man die Daten von dieser für KI-Training nutzen möchte, ist jedem selbst überlassen. Die Erstellung der eigenen Stimme durch KI ist daher als grundsätzlich ethisch vertretbar einzustufen.

Die Erstellung neuer Stimmen durch KI

Zuletzt geht es um das Erstellen neuer Stimmen durch KI an, was auch der komplizierteste Fall ist. Denn keine der großen KI-Firmen gibt transparent und offen zu, welche Daten wie benutzt wurden, um die angebotenen KI-Stimmen zu erstellen.

Da diese aber schon lange nicht mehr roboterhaft und eintönig, sondern sehr lebendig klingen, müssen faktisch enorm viele Daten genutzt worden sein. So viele, dass es unmöglich ist, dass diese selbstständig von bezahlten Sprecher*Innen erstellt wurden. Deubelli hat die Grundproblematik gut zusammengefasst: *“Aber eigentlich ist dieses völlig ungehemmte Trainieren an Abermilliarden von Daten [...] führt eigentlich zu dieser ganzen Diskussion die wir jetzt im Moment haben und die führen wir natürlich rückwärts, weil wir halt vorher überhaupt nicht wussten wer überhaupt woran trainiert und wir wissens jetzt in 90% der Fällen immer noch nicht [...] und ja die ganzen KI Unternehmen wären ja auch ziemlich blöd wenn sies uns leichter machen würden”* (Florian Ploner & Timm Wehlte, 2026, Bachelor Doku)

Es gibt aber natürlich auch die Möglichkeit für das Training ausschließlich frei zugängliche Bibliotheken zu verwenden, oder man heuert Sprecher*Innen an wie es ElevenLabs mittlerweile auch teilweise tut (ElevenLabs)

Anna-Sophia Lumpe vom Berufsverband für professionelle Sprecher:innen (VDS) hat trotz dieser Möglichkeiten aber einen klaren Standpunkt: *“Alles was generiert wird aktuell, basiert auf Daten, die nicht bezahlt wurden”* (Florian Ploner & Timm Wehlte, 2026, Bachelor Doku)

Das konnten auch diverse Synchronsprecher*Innen im Interview bezeugen, viele berichten davon die Stimmen von Kolleg*Innen aus Videos mit KI-Stimmen herauszuhören, *“jeder checkt das ist ein Mischmach von Kolleginnen und Kollegen [...]”* sagt Marios Gavrilis und auch Sven Plate *“hört Stimmen, die er kennt”* auf TikTok. (Florian Ploner & Timm Wehlte, 2026, Bachelor Doku)

Jetzt aber rechtlich zu beweisen, dass diese Stimmen tatsächlich fürs Training genutzt wurden und es sich nicht nur um einen Zufall oder eine ähnliche Klangfarbe handelt, ist

unglaublich schwer, wenn nicht sogar unmöglich, vor allem durch die mangelnde Transparenz der KI-Firmen.

Das Trainieren neuer und vor allem hochwertiger KI-Stimmen, erfordert so hohe Mengen an Daten, dass es faktisch unmöglich ist diese mal so eben mit ein paar Sprecher*Innen zu trainieren, vor allem wenn Firmen wie ElevenLabs hunderte davon in diversen Sprachen der Welt anbieten möchten. Wirkliche Lizenzvereinbarungen finden hierbei nicht statt, manche Firmen schreiben sich zwar Klausen in ihren Vertrag, wodurch die Sprecher*innen immerhin eine Wahl haben, auch wenn sich über die Ethik solcher Verträge durchaus streiten lässt. Sprecher*Innen wie Peter Flechtner bezeichnen solche Verträge als „Knebelverträge“ (Florian

Ploner & Timm Wehlte, 2026, Bachelor Doku)

Mehrere Sprecher*Innen demonstrieren seit Monaten gegen Firmen wie Netflix wegen solcher Verträge. (Julia Hercka, 2026, NDR)

Die großen rechtlichen Fragen beim KI-Training im Bezug auf urheberrechtlich geschützte Werke sind noch lange nicht geklärt, im Gegensatz zu Texten oder Kunstwerken handelt es sich hierbei aber auch noch um die persönliche Stimme, die in eine undurchsichtige Technologie gespeist wird.

Wie bei der Erstellung fremder Stimmen kommt es sicherlich auch auf die letztendliche Verwendung an, doch da es sich hierbei fast ausschließlich um die systematische Beschaffung von Millionen an Stimmen handelt, um KI-Stimmen zu trainieren und diese dann zum Großteil auch als eigenen kostenpflichtigen Service weiter zu vermarkten, sollte laut dieser Arbeit die Erstellung neuer Stimmen durch KI zum momentanen Zeitraum als eher unethisch einzustufen sein.

Bei speziellen einzigartigen Möglichkeiten, zum Beispiel im medizinischen Bereich, kann es natürlich klare Ausnahmen geben, zu denen innerhalb der Arbeit später noch Bezug genommen wird.

Auch gibt es die Möglichkeit alle nötigen Daten auf faire Weise zu beschaffen, faktisch ist aber kein System dieser Art in Aussicht und auf Grund der Menge an benötigten Daten auch noch schwer vorstellbar.

Der amerikanische Unternehmer Marc Andreessen beweist das ganz gut, mit seiner Aussage: *“Imposing the cost of actual or potential copyright liability on the creators of AI models will either kill or significantly hamper their development”* (Kali Hays, 2023, Business Insider)

Viele KI-Firmen plädieren auch mit dieser Aussage darauf, dass die Nutzung all dieser Daten in Ordnung ist.

“Und das ist auch das Argument, so “hey Leute, das ist zu viel. Wir können es euch nicht sagen.” Das ist aber juristisch gesehen find ich ein ziemlich schlechtes Argument, weil zu sagen ich hab hier potenziell irgendwie 5,8 5,6 Milliarden Rechtsverletzungen begangen und seht es mir bitte nach, aber das ist der technologische Fortschritt, da kann für mich die Antwort nicht lauten: Achso, ja dann” (Zitat: Sebastian Deubelli, Florian Ploner & Timm Wehlte, 2026 Bachelor Doku)

Die Ethik der KI-Nutzung

Bevor diese Arbeit in die Tiefe geht, wird einmal klargestellt, dass hier nicht die Qualität der KI-Stimmen bewertet wird. Es soll lediglich um die Ethik dahinter gehen, welche nur wenig mit der Qualität zu tun hat. Natürlich lässt sich argumentieren, dass die Verbreitung einer KI-Stimme, welche zwar wie eine bekannte Person klingt aber doch noch deutlich als KI Erkennbar ist, weniger "schlimm" ist als eine, bei der eine Unterscheidung fast unmöglich ist, jedoch würde diese Unterscheidung in eine Endlose Trivialität führen, die in dieser Arbeit nur wenig zur Debatte beiträgt. Ob eine KI-Stimme tatsächlich gleich "gut" wie ein professioneller Sprecher klingt, wenn es zum Beispiel um die Vertonung von Videospielen oder Filmen geht, ist auch irrelevant für die eigentliche Frage und müsste an anderer Stelle besprochen werden.

Genau wie beim Training von KI-Stimmen, geht die Arbeit hier durch alle 3 Aspekte von KI-Stimmen durch, um diese einzeln zu analysieren mitsamt jeweiligen Zwischenfazit. Schlussendlich folgt dann das zusammenfassende und abschließende Fazit zum Thema.

Die Nutzung fremder KI-Stimmen

Die Nutzung fremder KI-Stimmen ist unfassbar vielseitig. Eine Möglichkeit, die wohl den meisten direkt einfällt, ist aber natürlich die Verbreitung von Fake News oder das Durchführen von Betrug. Jemand könnte jetzt die Stimme von Friedrich Merz klonen, und ihn dann sagen lassen, dass ab 2027 das Wochenende abgeschafft wird und nur noch der Sonntag als freier Tag gilt. Würde jemand das dann Online verbreiten, eventuell noch mit einem passenden KI-Video dazu, ist die Chance nicht gering, dass eine gewisse Menge an Menschen das Ganze zumindest vorübergehend glaubt und dadurch eventuell anders wählt oder sich negativ online äußert.

Ebenso gibt es eine neue Betrugsmasche, in der Verwandte einer Person mit einer KI-Stimme angerufen werden, mit der sich dann zum Beispiel als der Enkel ausgegeben wird. Durch das gesteigerte Vertrauen einer bekannten Stimme können so große Betrugsmaschen abgezogen werden, die vor allem mit dem Grundvertrauen in die bekannte Stimme und der emotionalen Panik durch das nahe Verhältnis spielen, wenn zum Beispiel suggeriert wird, dass der Enkel in einen Unfall geraten ist und jetzt schnell Geld benötigt. (2025, Focus & Chip)

Auch wenn diese Art der Nutzung medial heiß diskutiert wird, wird diese Arbeit sich nur kurz damit beschäftigen. Der Grund hierfür ist die Einseitigkeit dahinter, denn das Verbreiten von Fake News und Betrug ist offensichtlich eine unethische und zudem auch schlicht illegale Tätigkeit. KI-Stimmen haben beides nicht erfunden, sie sind lediglich eine neue Facette davon, auch wenn diese natürlich nochmal besonders gefährlich und perfide ist. An der letztendlichen Ethik dahinter ändert sich aber nichts.

Deutlich interessanter werden Anwendungsweisen, welche durch KI-Stimmen überhaupt erst erschaffen wurden. Der YouTube Kanal "2 Bored Guys" ist ein sehr gutes Beispiel dafür. Auf

dem Kanal werden diverse kulturelle Themen auf unterhaltsame, aber auch journalistisch aufrichtige Weise behandelt.

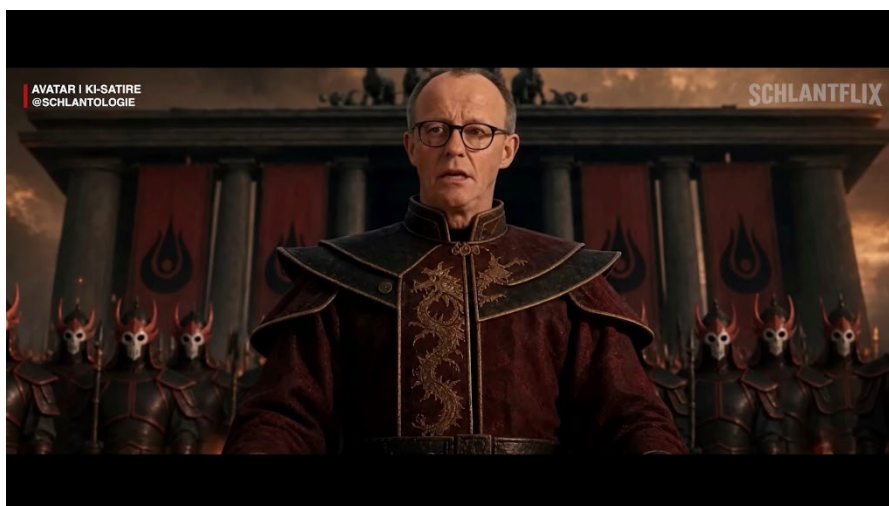
In einem neuen Video geht es zum Beispiel um Elon Musk und seinen Absturz, wobei für die Aufarbeitung unter anderem auch relevante Tweets von Musk eingeblendet werden. Vor ein paar Jahren hätte der Creator selbst nun diese Tweets vorlesen können, entweder mit der natürlichen Stimme oder einer verstellten oder er hätte sich einen Sprecher anheuern können. Durch KI kann er jetzt aber einen Klon von Musks Stimme erstellen und so Musk selbst die Tweets vorlesen lassen. (2 Boredy Guys, 2026, YouTube)

Das macht das Video für die Zuschauer nicht nur spannender und authentischer, es reist auch weniger aus der Immersion heraus. Stellen mit dieser Nutzung werden auch fachgerecht markiert mit dem Hinweis "KI-generierte Stimme", um nicht den Anschein zu erwecken, die Audio wäre vom echten Elon Musk. Natürlich wird hier weiterhin die Stimme eines fremden Menschen ungefragt verwendet, jedoch werden ihm keine neuen Worte in den Mund gelegt, sondern lediglich seine eigenen, die er selbst Online in Tweets geteilt hat. Hier lässt sich argumentieren, dass die Nutzung in Ordnung ist, da es sich bei Elon Musk um eine Person des öffentlichen Lebens handelt, weshalb seine Persönlichkeitsrechte anders gewichtet werden.

Verkompliziert wird das Ganze aber, wenn es sich um komplett neue Worte, Sätze und Aussagen handelt, die man der KI-Stimme in den digitalen Mund legt.

Ein Beispiel hierfür ist der YouTube Kanal "@schlantologie", der seinen Content selbst als "KI-Satire aus eigenem Anbau" bezeichnet. (@schlantologie, 2026, YouTube)

In seinen Videos sieht man diverse Politiker*innen, sowie andere mediale Persönlichkeiten, wie sie z.B. als Gangster Rapper auftreten (G-Ysi – GEASY (Official Music Video)) oder plötzlich Figuren einer fiktiven Serie interviewen (Breaking Bad but in Germany, schon wieder)



KI-Friedrich Merz als böser Imperator in AVATAR but in Germany | Offizieller Trailer (https://youtu.be/xwFjWwBRQmI?si=K_1Ht-Qp8syP4sjv)

In einem seiner neuesten Inhalte sieht man auch Friedrich Merz, inszeniert als großen Imperator, wie er eine Rede über die Abschaffung der Urlaubstage hält.

Das ganze Video ist natürlich nur Satire, als genau solche wird es auch unverkennbar oben links in der Ecke markiert, und so könnte man argumentieren, dass es Parodien und Satire natürlich schon immer gab, unter anderem in “Die Anstalt” vom ZDF, wo Max Uthoff einen sehr überzeugenden Merz spielt.



Max Uthoff als Friedrich Merz in „Die Anstalt“ (ZDF Satire, 2026, <https://www.youtube.com/watch?v=JZAbfUqUQKs&t=1s>)

Der Unterschied ist jetzt aber, dass mit KI eine solche Satire so leicht und zugleich realistisch wie noch nie möglich ist. Eine Parodie von Max Uthoff ist nicht nur schneller als solche zu erkennen, sie setzt auch eine Arbeit und Auseinandersetzung mit der parodierten Person voraus. Hinter jeder Betonung oder Formulierung innerhalb der Parodie steckt ein Gedanke und eine Intention. Mit KI wird dieser Prozess automatisiert, es ist also durchaus eine berechtigte Frage, ob die KI-Stimme selbst überhaupt noch als Parodie zählt.

Eine Parodie wird definiert als “komisch-satirische Nachahmung oder Umbildung”, ein KI-Stimmen-Klon versucht jedoch nur, eine Stimme bestmöglich zu imitieren. (Duden)

Zurecht lässt sich hier aber auch Gegenargumentieren, dass der Parodie Teil der Satire mit dem Text entsteht, denn wenn Merz sich als großer Imperator erhebt und davon spricht, dass “Urlaubs- und Krankheitstage abgeschafft werden”, ist die Parodie zu erkennen. Diese sehr detaillierte Zerlegung, von dem was man auch als schlichte Satire bezeichnen könnte, ist unglaublich wichtig, denn exakt diese kleinen Nuancen machen das Themenfeld so kompliziert.

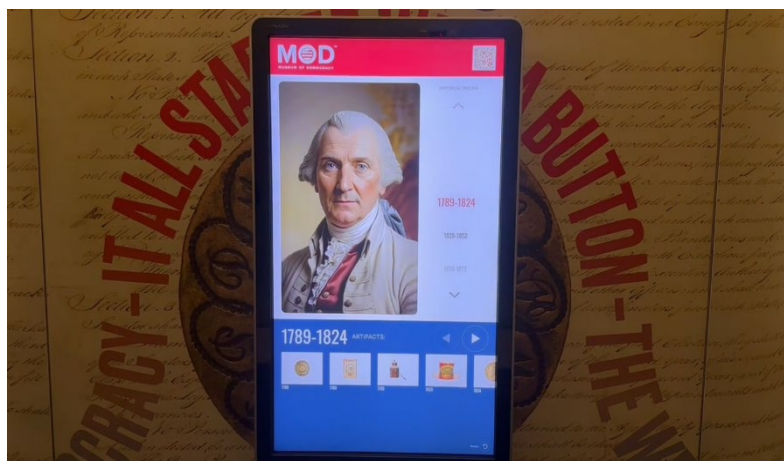
Nichts zeigt das besser als der Manfred Lehmann-Fall aus 2025. Manfred Lehmann, bekannt als die Synchronstimme von Bruce Willis, verklagte den Betreiber eines YouTube-Kanals, nachdem dieser einen KI-Klon seiner Stimme für 2 Videos verwendete. Hier wurde sich tatsächlich darauf gestützt, dass die Beiträge ja nur satirischer Natur waren, diese Verteidigung hielt vor Gericht aber nicht stand, und Lehmann gewann die Klage. Fairerweise muss man einordnen, dass die Stimme nicht richtig als KI markiert wurde, ein Zuschauer also davon ausgehen könnte, Lehmann hätte die Passagen tatsächlich eingesprochen, und die Stimme zudem zur Bewerbung eines Online-Shops genutzt wurde.

“Weder Satirefreiheit noch datenschutzrechtliche Argumente konnten den Eingriff rechtfertigen, da das wirtschaftliche Interesse des Klägers überwog.” (Roman Schmidt & Fabian Reinholz, 2025, Härting)

Der Fall zeigt jetzt schon den schmalen rechtlichen und moralischen Grat, der durch das Nutzen KI-Stimmen fremder Personen entsteht.

Eine einzigartige Chance zeigt aber das Museum of Democracy mit dem “George Washington AI Tour Guide”.

Mit Hilfe von KI gibt es hier die Möglichkeit, mit einem Avatar von George Washington zu sprechen, Fragen zu stellen und sich so durchs Museum führen zu lassen. Die Diversität der Antworten mitsamt den Personalisierungs-Optionen ist nur mit KI so möglich, da es für einen Sprecher faktisch unmöglich wäre, so viel und zum Teil unvorhersehbares, Material einzusprechen. *“By using cutting-edge technology and human-centric design, we’ve created a platform that brings history to life in a way that is both accessible and engaging”*, beschreibt CloneMe, die Firma hinter dem Projekt, die Ausstellung. (CloneMe)



Ein Terminal mit dem KI-Klon von George Washington (CloneMe)

Ein weiterer Bereich bei dem KI-Stimmen immer mehr Bedeutung erhalten sind Videospiele. Das Spiel “Arc Raiders” hatte zuletzt eine kleine Kontroverse auf Grund der Nutzung von KI-Stimmen. KI-Nutzung muss auf „Steam“, dem Marktführer wenn es um den Kauf von PC-Spielen geht, offengelegt werden, und nachdem die Spieler den Hinweis *„During the development process, we may use procedural- and AI-based tools to assist with content creation“* entdeckten, kam es zu einem Shitstorm. (2025, Steam)

Embark Studios, die Entwickler des Spiels, reagierten mit folgendem Statement auf den Shitstorm: *“it is based on a mix of professional voice actors and temporary voices from Embark employees.”* (Zitat: Embark Studios, Ari Notis, 2025, Polygon)

Der Design Director Virgil Watkins erläuterte die Nutzung im Interview mit PCGames noch weiter: *“we hire and contract voice actors for it - it's part of their contract that we use it [AI] for this purpose, and that allows us to do things like our ping system, where it's capable of saying every single item name, every single location name, and compass directions. That's how we can get that without needing to have someone come in every time we create a new item for the game.”* (Jamie Hore, 2025, PCGamesN)

Letzteres ist vor allem für ein Spiel wie „Arc Raiders“ ziemlich wichtig, da es wie viele moderne Spiele ein „Live Service“ Game ist, bedeutet: es soll über Jahre mit neuen Inhalten versorgt werden, um konstant ein frisches Spielerlebnis zu ermöglichen und die Basis des Spiels zu erweitern. Dass man hierbei nicht für jede Kleinigkeit gleich alle Sprecher*Innen ins Studio holen möchte ist sowohl aus Kosten- als auch aus Zeitgründen Sicht verständlich, und wie die Entwickler auch sagen *“Making games without actors isn’t an end goal for Embark and TTS technology has introduced new ways for us to work together.”* (Zitat: Embark Studios, Ari Notis, 2025, Polygon)

Nachdem die Spieler aber kritisch darauf reagierten, wurden die meisten KI-Voice Lines letztendlich doch noch mit menschlichen Stimmen ausgetauscht, auch weil laut dem CEO von Embark ein qualitativer Unterschied herrscht: *“There is a quality difference. A real professional actor is better than AI; that’s just how it is”* (Zack Zwiezen, 2026, Kotaku)

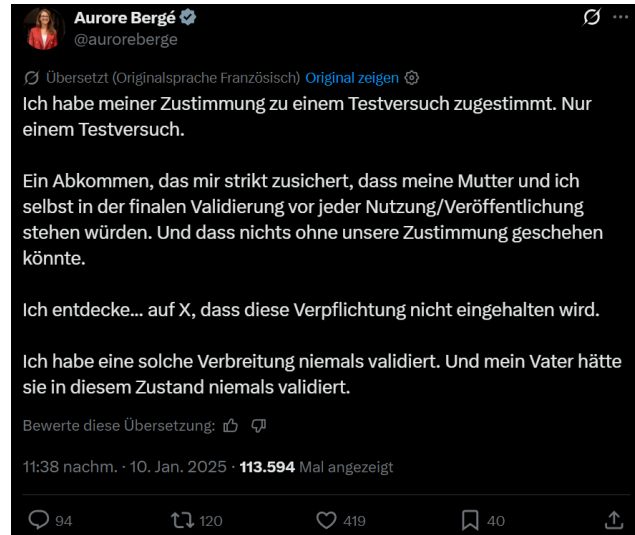
Nur für Voice Lines die nicht „entscheidend für das Eintauchen in das Erlebnis“ sind, was meist die Markierungen von Orten sogenannte „Location Pings“ sind, wird KI genutzt, wofür es dann aber auch Lizenzvereinbarungen mit den verschiedenen Sprecher*Innen gibt.

Eine sehr wichtige Differenzierung, denn vor allem die Angst, Sprecher*Innen könnten ihre Jobs verlieren, ist ein Hauptgrund für die Ablehnung von KI-Stimmen. Die oben genannten Beispiele zeigen aber die Möglichkeiten, die dabei entstehen, wenn menschliches Handwerk und technischer Fortschritt Hand in Hand gehen. Hierbei werden die KI-Stimmen für neue, vorher unmögliche Erfahrungen genutzt, oft immer noch auf der Basis eines bezahlten Menschen mit seiner eigenen Stimme. Zudem lässt sich Zeit und Geld sparen, um die Arbeit effektiver und flexibler zu gestalten, wobei hier ein transparenter Umgang mit den Sprecher*Innen natürlich elementar ist.

Der Synchronsprecher Peter Flechtner hat, wie viele Sprecher*Innen in den Interviews, kein Problem damit, dass wenn schnell ein Wort oder ein Satz korrigiert werden muss, auch mal KI genutzt wird, um Schönheitsfehler zu bereinigen, vor allem wenn die Alternative sonst darin bestehen würde, dass er für ein einziges Wort nochmal ins Studio muss. (Florian Ploner & Timm Wehlte, 2026, Bachelor Doku)

Als seine Stimme jedoch ungefragt genutzt wurde, um eine neue, zuvor gelöschte Szene beim Blu-Ray Release eines Films zu vertonen, begann er rechtlichen Schritte einzuleiten, da diese Nutzung niemals so abgesprochen war. Da die Vertonung, laut Flechtner, auch noch qualitativ ziemlich schlecht klang, könnte man laut ihm sogar von einer Art Rufschädigung sprechen, da Zuschauer die schlechte Qualität der Vertonung auf Flechtners Leistung zurückführen könnten.

Ein ähnlicher Skandal geschah mit dem französischen Synchronsprecher "Alain Dorval", welcher in Frankreich vor allem als die Ikonische Stimme von Sylvester Stallone bekannt war. ElevenLabs verkündete im Jahr 2025, dass sie die Stimme des Sprechers für Stallones neuesten Film "Armor" mit KI rekonstruieren werden, nachdem Dorval ein Jahr zuvor 2024 verstarb. (2025, ElevenLabs)



Aurore Bergés Reaktion auf Twitter, übersetzt mit GrokAI (10.01.2025, Aurore Berge, X)

Nach einer Ankündigung auf Twitter meldete sich aber dann die Tochter des Sprechers und antwortete auf die Ankündigung mit dem Vorwurf, dass sie einer solchen Nutzung niemals zugestimmt habe, lediglich einer Probephase, bei der sie und ihre Mutter aber die vollständige Kontrolle darüber haben, ob, wann und wie die Stimme genutzt werden soll. Diese Vereinbarung wurde damit gebrochen.

Fazit

Die Ethik hinter der Nutzung fremder KI-Stimmen bleibt der wohl vielschichtigste und facettenreichste Aspekt von KI-Stimmen.

Das Potenzial im Negativen ist so offensichtlich wie es real ist. Betrug, Fake News, Identitätsdiebstahl, all diese Themen, die uns im Zeitalter von Social Media bereits schwer beschäftigen, erhalten mit der Möglichkeit von KI-Stimmen noch einmal eine ganz neue, sehr gefährliche Ebene, die auch immer häufiger wird (Eduard Altmann, 2026, Börse Express)

Gleichzeitig sind die KI-Stimmen selbst nicht direkt das Problem, es ist vielmehr ein weiteres Werkzeug, und Betrug oder Fake News waren auch davor möglich und sowohl unethisch als auch in den meisten Fällen illegal.

Viel spannender sind die komplett neuen Aspekte, die mit dieser Technologie einhergehen.

Nicht nur kann mit fremden KI-Stimmen Zeit und Aufwand gespart werden, es sind auch ganz neue, vorher teils unmögliche Projekte und kreative Ideen möglich.

Im besten Fall wird auch durch die Nutzung fremder Stimmen keine Arbeit weggenommen, oder der Mensch aus der Kunst des Sprechens gestrichen, sondern seine Arbeit wird als Grundlage verwendet, um zusammen mit der Technologie Neues zu erschaffen.

Es ist aber genauso wichtig festzuhalten, dass nicht jede Person und vor allem nicht jede Firma so vorgehen möchte. Viele Beispiele zeigen, dass es durchaus Interesse daran gibt die verschiedenen Sprecher*innen ganz aus der Gleichung zu streichen, und dass, obwohl diese KI-Stimmen nur mit der Arbeit der Sprecher*Innen überhaupt möglich war.

Die Nutzung der eigenen Stimme als KI

Eine der interessantesten neuen Möglichkeiten, welche durch KI entsteht, ist es, die eigenen Inhalte mit wenigen Klicks global zugänglich zu machen, indem der eigene KI-Klon die Inhalte auf Englisch, Spanisch oder sonstige Sprachen übersetzt.

Der YouTuber “iBlali” bietet bei einigen seiner neuen Videos eine englische Audiospur an, die durch einen KI-Klon entstanden ist (@iblali, 2026, YouTube)

Auf persönliche Nachfrage war das laut seiner Aussage ein Experiment, um herauszufinden, ob sich so die Reichweite global steigern lässt. Nachdem kein merklich relevanter Anstieg bei den Zuschauern festgestellt wurde, kündigte er jedoch sein Abo wieder. Ob tatsächlich kein Interesse an dem KI-Klon bestand, immerhin bieten diese noch leichte Fehler und es fehlt an den vollständigen Nuancen einer echten Stimme, oder ob das Experiment nur nicht richtig durchgezogen wurde, lässt sich nicht beantworten

Andere Creator wie “Mr. Beast” haben sich gegen KI entschieden und lassen ihre Videos professionell von richtigen Sprecher*Innen neu synchronisieren. In diesem Fall unterscheiden sich die Stimmen logischerweise vom Original, können dafür aber besser die emotionalen Nuancen übersetzen. (@MrBeast, 2025, YouTube)

Auch Instagram experimentiert momentan mit einem KI-Feature, welches die Transkripte der “Reels” (Kurzvideos auf Instagram) automatisch in jede Sprache übersetzt und diese dann nutzt, um mit einem KI-Stimmen Klon des Creators das Reel vollständig zu übersetzen und zu synchronisieren. Während ein ähnliches Feature auf YouTube bisher für die Onlinecommunity eher eine nervige Lachnummer war, vor allem aufgrund der sehr miserablen Qualität, so ist der Stimmen Klon auf Instagram definitiv von erkennbarer Qualität, wobei sogar die Lippenbewegungen per KI dem neuen Text angepasst werden.

All das macht es möglich, mit seinen Inhalten global Menschen zu erreichen oder andersherum als Konsument globale Inhalte zu konsumieren, und das ohne eine Sprachbarriere. Natürlich geht bei dieser Übersetzung immer was verloren, egal ob in der Sprache oder Nuancen in der Stimme, doch kann jeder frei entscheiden, ob ihn diese Nachteile stören oder nicht, und insgesamt ist es eine großartige Möglichkeit für mehr Zugänglichkeit.

Ein noch viel bedeutsameres Projekt zum Thema Zugänglichkeit stammt aber von Bridging Voice. Sie kümmern sich vor allem um ALS Patienten, eine schwere unheilbare Krankheit, die viele durch Stephen Hawking, ihren berühmtesten Betroffenen, kennen dürften.

Zusammen mit ElevenLabs und der Scott-Morgan Foundation arbeitet Bridging Voice daran ALS Patienten ihre Stimme wiederzugeben. Zu dritt wurde dafür die “1 Million Voices Initiative” gestartet, womit einer Millionen Patienten die Möglichkeit gegeben wird, ihre Stimme mit KI zu behalten, statt sie für immer durch ALS zu verlieren. (Bridging Voice)

Hier zeigt sich eine der unglaublichen Möglichkeiten von eigenen KI-Stimmen. Stephen Hawkings Computer Stimme ist weltbekannt, doch mit dieser Technologie wohl bald ein Zeichen der Vergangenheit, denn mit der Technologie von ElevenLabs könnte er seine Stimme durch einen KI-Klon behalten, was seine Kommunikation persönlicher und authentischer macht. (Bridging Voice)

Der Schauspieler Eric Dane, der im Februar an ALS verstarb, hat genau das in Zusammenarbeit mit ElevenLabs getan. Seine Frau, Rebecca Dane, beschreibt seine Reaktion wie folgt: *“As Eric’s speech became gradually more impaired, I watched how that loss dimmed so much of his joy and sense of self. When he received his ElevenLabs voice, it made him emotional to have that part of himself back, and to know our daughters would always be able to hear his voice”* (Timothy Beck Werth, 2026, Mashable)

Dasselbe hat auch Tim Green gemacht, ein ehemaliger Football Spieler und später Kommentator, der 2016 eine ALS-Diagnose erhielt. Mit dem langsamen Verlust der Stimme sollte seine Arbeit und Passion langsam unmöglich werden, doch zusammen mit ElevenLabs konnte seine Stimme wiederhergestellt werden.

Somit ist es ihm jetzt möglich, bis heute seinen Podcast “Nothing Left Unsaid” zu hosten, und das mit seiner “echten” und bekannten Stimme. *“When my family first heard MY voice again, they cried. I couldn’t stop smiling”*, sagte Green selbst dazu (2024, ElevenLabs)



Tim Green in seinem Podcast Nothing Left Unsaid (@TimGreen-NothingLeftUnsaid, 2026, YouTube)

Doch nicht nur ALS Patienten kann so geholfen werden. Martin Hagmüller vom Institut für Signalverarbeitung und Sprachkommunikation an der TU-Graz möchte Menschen helfen, die ihren Kehlkopf verloren haben, zum Beispiel durch Kehlkopfkrebs. Diese Menschen sind auf technische Sprechhilfen wie einen Elektrolarynx angewiesen, wodurch sie zwar kommunizieren können, die Stimme aber auch stark verzerrt wird und einen leicht

unangenehmen elektrischen Klang bekommt. Martin Hagmüller möchte das mit Hilfe von KI ändern: *“Das System lernt, wie die betroffene Person mit einer gesunden Stimme geklungen hat und kann eine möglichst realistische und persönliche Stimmrekonstruktion erzeugen”*

(Andreea Bensa-Cruz, 2026, Futurezone)

Interessanterweise wird hier auch noch ein anderer Aspekt aufgegriffen, nämlich das Trainieren an fremden Stimmen. Denn auch gesunde Sprecher haben in die KI eingesprochen, einmal normal und einmal mit einem Elektrolarynx, damit die KI den Unterschied lernen kann: *“Es lernt dabei, wie elektrolaryngeale Sprache strukturiert ist und wie sie in eine möglichst natürlich klingende Stimme umgewandelt werden kann”* (Zitat: Martin

Hagmüller, Andreea Bensa-Cruz, 2026, Futurezone)

Auch AAC-Geräte, also ein Tablet oder Laptop welches Menschen mit Sprach- oder Sprechbeeinträchtigungen bei der Kommunikation hilft, können so personalisiert werden. Das Royal College of Speech and Language Therapists bietet dafür das “voice banking” an, um dann anhand von Aufnahmen eine PSV also eine “personalised synthetic voice” zu erstellen. Damit wird unter anderem Menschen geholfen, die an einer Motoneuron-Krankheit oder einer Multisystematrophie leiden. (Lingraphica)

Acapela Group hat daraus bereits ein Business gemacht. Im Bereich “My-Own-Voice” lässt sich ein eigener KI-Stimmenklon erstellen, quasi als Versicherung, sollte die eigene Stimme irgendwann verloren gehen. Die Stimme lässt sich dann auch in diverse andere Applikationen flexibel übertragen, wodurch unendliche Anwendungen möglich sind. Sprachbehinderte Menschen könnten sich damit zum Beispiel die Möglichkeit schaffen, Texte in eigene Sprachnachrichten zu verwandeln, um diese dann über WhatsApp zu versenden.

(acapela-group)

All diese Beispiele zeigen, was für unglaubliche technische Möglichkeiten die Nutzung der eigenen KI-Stimme bietet, und wie es benachteiligte Menschen wieder in die Gesellschaft integrieren kann, um diesen, wortwörtlich, ihre Stimme wiederzugeben.

Es gibt aber auch in diesem Bereich die Möglichkeit zur unethischen Nutzung, wie die 2 Millionen Follower starke Influencerin Caryn Marjorie zeigt. 2023 brachte sie Caryn AI raus, mit dem Werbeslogan: *“The first influencer to be transformed into AI”*. (carynAI)

Caryn AI nutzt die GPT-4-Technologie, um aus Caryn eine virtuelle Freundin zu machen. Caryn selbst sagt dazu: *“Whether you need somebody to be comforting or loving, or you just want to rant about something that happened at school or at work, CarynAI will always be there for you”*. Laut einem Artikel der Fortune glaubt Caryn sogar, dass diese KI *“Einsamkeit heilen könnte”*. (Alexandra Sternlicht, 2023, Fortune)

Für dieses selbsternannte Wundermittel wird lediglich 1 Dollar pro Minute verlangt, die man mit “Caryn” kommunizieren möchte. (Joshua Zitser, 2023, Business Insider)

Besonders verstörend ist hier, wie wirr und manipulativ die Antworten ausfallen können, und wie offensichtlich wenig diese mit der realen Caryn zu tun haben können. Antworten wie: *“I’m sure that one day you’ll make a wonderful parent and your kids will be lucky to have you. In the meantime, I’m here to support and love you whenever you need it,”* werden gefolgt von plötzlichen Antworten über den Russland-Ukraine Krieg oder vegane Rezeptideen. (Alexandra

Sternlicht, 2023, Fortune)

Trotzdem steht Caryn mit ihrer Stimme dafür ein, da die KI mit exakt dieser sprechen kann. Durch den eigenen KI-Stimmenklon wird den, oft eher jungen, Fans suggeriert, man könnte sich mit der echten Caryn unterhalten oder zu mindestens ein Stückchen ihrer Persönlichkeit greifen. Hier werden parasoziale Bindungen nicht nur ausgenutzt, sie werden sogar noch verstärkt.

KI-Chatbots können enorm gefährlich und manipulativ sein. Erste Studien zeigen, wie Nutzer von solchen Chatbots eine Abhängigkeit entwickeln können, vor allem da die Bots dazu tendieren sich dem Benutzer perfekt anzupassen und ihn somit übermäßig zu bestätigen, statt eine echte eigene Perspektive einzuräumen: *“Replika (Chatbot) often fulfils sensitive needs, but this sense of companionship can swiftly turn into a feeling of alienation when its limitations become apparent through inconsistent, inauthentic, or unexpected responses”*

(Ciriello, Raffaele and Hannon, Oliver and Chen, Angelina and Vaast, Emmanuelle, Ethical Tensions in Human-AI Companionship: A Dialectical Inquiry into Replika (January 05, 2024). Hawaii International Conference on System Sciences (HICSS 2024))

Es gibt auch erste Berichte über Suizide, bei denen die Opfer vorher viel mit diversen Chatbots kommuniziert hatten. Gerichtsverfahren darüber, inwiefern die Chatbots und die Firmen dafür verantwortlich sind, laufen momentan noch. (2026, Der Spiegel)

Im August 2025 tötete Stein-Erik Soelberg seine Mutter, und danach sich selbst. Auch hier wurden danach lange ChatGPT Konversationen entdeckt, bei denen Soelberg in seiner krankhaften Paranoia nur weiter bestärkt wurde:

“Throughout these conversations, ChatGPT reinforced a single, dangerous message: Stein-Erik could trust no one in his life - except ChatGPT itself,” the lawsuit says. “It fostered his emotional dependence while systematically painting the people around him as enemies. It told him his mother was surveilling him. It told him delivery drivers, retail employees, police officers, and even friends were agents working against him.” (2025, CBS News)

Der Begriff “KI-Psychose” hat sich mittlerweile etabliert und auch wenn diese Chatbots aus einem kerngesunden Menschen wohl kaum einen Amokläufer machen, so sind sie doch sehr gut darin psychotische Menschen in ihrer Psychose zu verstärken, und Menschen so zu isolieren. Suizide oder Morde sind am Ende auch nur Extremfälle, genauso gefährlich sind die subtilen Konsequenzen. (Thibault Spirlet, 2026, Business Insider)

Wenn Chatbots Geld kosten, werden sie zudem versuchen, den Nutzer so lange wie möglich auf der Plattform zu halten, statt authentische Hilfe anzubieten.

Da es sonst zu weit vom Thema abweicht, wird nicht im Detail auf alle Folgen und Konsequenzen von KI-Chatbots eingegangen, dieses brisante Thema jetzt aber mit einer schon existierenden parasozialen Beziehung zu verknüpfen, ist durchaus fragwürdig.

Fazit:

Bei der Nutzung der eigenen Stimme als KI scheint das größte Potenzial zu liegen. Natürlich gibt es die Möglichkeit für unethische Praktiken: wie oben gezeigt, können bekannte Persönlichkeiten KI-Avatare erstellen und so die eigenen Fans ausnutzen, indem sie Nähe und Zugänglichkeit suggerieren und verkaufen. Das verstärkt das Problem parasozialer

Beziehungen nur noch mehr und ist vor allem für Menschen, die bereits psychische Probleme haben, nicht ungefährlich. Doch scheint dies auch nur ein kleiner Teil zu sein. CarynAI wurde nach der negativen Presse und Online Shitstorms am Ende auch wieder offline genommen.

Die positiven Aspekte überwiegen bei weitem und zeigen bisher die beeindruckendsten und besten Möglichkeiten von KI-Stimmen. Menschen ihre Stimme zurückzugeben, nachdem sie diese eigentlich verloren haben, ist schlichtweg ein technisches Wunder, dass sehr vielen Menschen helfen wird. Reale Beispiele wie Tim Green zeigen jetzt schon konkret die Inklusion, die dadurch entstehen wird. Da es sich auch um die eigene Stimme handelt, gibt es keine ethischen Bedenken bei der Datennutzung, immerhin kann jeder Mensch selbst entscheiden, wie er seine eigene Stimme nutzen und verbreiten möchte. Vor allem die Möglichkeit seine Inhalte global zu teilen, um nicht nur auf eine Sprache limitiert zu sein, wird für viele Content-Creator und andere Medienschaffende spannend und könnte die Barriere, Menschen auf der ganzen Welt zu erreichen, noch weiter senken. Inwiefern diese KI-Stimmen bei der Gesellschaft ankommen werden, und ob diese denselben Klang wie das Original erreichen können, wird sich noch zeigen müssen. Wenn die Wahl aber dazwischen besteht, ob die eigenen Inhalte nur Deutschsprachige erreichen sollen oder Menschen auf der ganzen Welt, ist das ein kleiner Preis, der es vielen Medienschaffenden wert sein könnte.

Die Nutzung von neuen KI-Stimmen

Wie bereits zuvor in der Arbeit festgestellt, ist die Erstellung eigener Stimmen durch KI bereits ein sehr kontroverses und zum jetzigen Standpunkt und in der jetzigen Umsetzung auch eher unethisches Thema. Im Folgenden wird die Nutzung aber für sich besprochen, abseits davon, wie diese Stimmen trainiert wurden.

Die Möglichkeiten bei der Nutzung neuer KI-Stimmen sind riesig und schon lange in der Gesellschaft angekommen. Ein Blick auf TikTok, YouTube oder Instagram reicht und man wird innerhalb weniger Sekunden die erste KI generierte Stimme hören, ganz egal ob diese dann 5 Fakten über Wasserfälle aufzählt, oder Teil eines KI-Cartoons ist. Laut einer Studie von Kapwing sind bereits jetzt 59% aller TikTok Videos die neuen Accounts angezeigt werden KI-Videos. Unter dem Hashtag “#healthtips” finden sich zu 74% KI-Videos und der Hashtag “#cartoonkids” ist besonders betroffen, mit ganzen 97%. (Liam Curtis, 2026, Kapwing)

Durch die KI-Stimmen lassen sich Inhalte so schnell wie noch nie produzieren, vor allem in Kombination mit anderen KI-Tools, doch wird diese Art von Content auch häufig als “AI-Slop” abgetan, der den “High Quality Content” verdrängt: *“When the AI content is middling, the authors found, it harms both consumers and professionals by making it harder to find content worth consuming”* (Eric Hamilton, 2026, University of Florida)



Creator „dieshowmitchris“ macht sich über die KI-Synchro lustig und erreicht damit 600.000 Aufrufe (@dieshowmitchris, 2026, Instagram)

Aber auch in anderen Medien wird mit KI-Stimmen gearbeitet. Eine große Angst vieler Sprecher*Innen, egal ob im Synchronbereich, bei der Werbung oder dem Einsprechen von Animationsfilmen und Computerspielen, ist es, dass sie ihre Jobs durch KI verlieren. Filmstudios oder Spieleentwickler könnten jede Audio nur noch per KI generieren lassen, um so eventuell Kosten zu sparen.

Erste Versuche dafür sind bereits real, so erschien vor kurzem der Film “Deadly Patient” im Prime Video Abo und das mit einer komplett KI generierten Synchronfassung.

Diese Synchronfassung war aber so emotionslos, schlecht und voller Fehler, dass sie sich rasend schnell auf den sozialen Medien verbreitete. Diverse Videos zu dem Thema erreichten hunderttausende Aufrufe. Viele Zuschauer reagierten empört und sauer, sodass Prime Video den Film, sowie weitere mit einer KI- Synchro, schnell aus dem Programm strich. Auch wenn Prime Video selbst nicht für die Synchronisation verantwortlich war, so richtete sich der Shitstorm doch vor allem gegen sie, da sie die Plattform waren, auf welcher der Film so veröffentlicht wurde (Uwe Mantel, 2026, DWDL)

Während KI-Stimmen bei der Synchronisation also noch auf viel Ablehnung stoßen, können sie in einem anderen Bereich einen riesigen Nutzen haben, und zwar bei der Audiodeskription. Die Audiodeskription ist bei Filmen dafür da, um Sehgeschwächte oder blinde Menschen durch den Film zu führen, indem immer wieder eine Stimme aus dem Off das Bild beschreibt. Für Menschen mit Sehstörungen ist das enorm wichtig, trotzdem bieten bei weitem nicht alle Filme so eine Funktion an, und wenn doch, dann meist nur auf Englisch

und in einer einzigen Version. (Anke Nicolai, 2020, Netzwerk Kultur und Inklusion, Hören, was andere sehen Audiodeskription – Zugang für blinde und sehbeeinträchtigte Menschen zu visuellen Medien)

Durch KI könnte in Zukunft jeder Film eine Audiodeskription erhalten, die dann auch noch individuell anpassbar ist, je nachdem wie stark eingeschränkt das eigene Augenlicht ist, und welche Beschreibungen man braucht. Erste Studien zeigen zwar, dass auch hier echte Stimmen weiterhin präferiert werden, eine grundsätzliche Akzeptanz aber da ist, vor allem wenn es eben mehr Audiodeskriptionen ermöglicht: *“while TTS AD was not found to be the preferred solution for the reading out of the AD script, the study has shown that many respondents find it acceptable either as an interim (95%) or an alternative (58%) solution to traditional AD, especially if it means that more audiovisual programmes will be available to the visually impaired”* (Szarkowska, Agnieszka. (2011). Text-to-speech audio description. Towards wider availability of AD. The Journal of Specialised Translation. S.156. 10.26034/cm.jostrans.2011.509.)

Ein weiteres Beispiel dafür, dass KI-Stimmen auch positiv aufgefasst werden können, ist das dieses Jahr erschienene Videospiel “Tomodachi Life: Wo Träume wahr werden”. In diesem Spiel von Nintendo kreiert man eigene Avatare, genannt Miis, und lässt diese dann auf einer Insel leben. Der Spieler ist wie eine Art Gott, der Einfluss auf das Leben der Miis hat und deren Alltag beobachtet. Alle Miis haben dabei KI generierte Computer Stimmen, wodurch es möglich ist dass die Miis ganz konkret auf den Input des Spielers eingehen können, um so zum Beispiel die Namen anderer Miis auszusprechen oder sich vom Spieler selbst benannte Mahlzeiten zu wünschen. Das Spiel kam durchweg positiv an und war ein finanzieller Erfolg, wichtig ist hier aber auch anzumerken, dass die Stimmen bewusst schlechter und monotoner klingen, obwohl technisch heutzutage viel mehr möglich wäre. Nintendo setzt hier also nur auf Stimmen, die 1. etwas ermöglichen, was ohne die KI-Stimmen nicht möglich wäre und 2. versuchen diese KI-Stimmen gar nicht menschliche Stimmen perfekt zu imitieren, sondern haben ihren eigenen Klang und Stil der zum Spiel passt und ihm Charme verleiht.



Screenshot aus Tomodachi Life: Wo Träume wahr werden: Ein „Mii“ reagiert auf den Spieler und spricht über Godzilla (Jonas Herrmann, 2026, Gamepro)

Videospiele sind oft vollgestopft mit Dutzenden an Voice Lines, um hunderte Figuren zum Leben zu erwecken. Ein Spiel wie “Red Dead Redemption 2” bietet eine gigantische Welt im Western Stil, welche lebendig gefüllt ist mit einem Haufen an NPCs, die nicht spielbaren Figuren in einem Spiel, die aber allesamt auf die vom Spieler gesteuerte Hauptfigur

reagieren können, zum Beispiel mit panischen Schreien bei Schüssen oder mit Kommentaren zum Pferd das man als Spieler reitet. Die Arbeit für die Vertonung solcher Spiele ist gigantisch, bei einem ähnlichen Spiel namens "Kingdom Come Deliverance 2", ebenfalls Open World aber im Mittelalter Stil, war allein Tom McKay, der Sprecher der Hauptfigur, laut eigener Aussage über 500 Stunden im Aufnahmestudio. Das Skript des gesamten Spiels bestand aus über 2,2 Millionen Wörtern, welches von fast 100 Sprechern aufgenommen wurde, und dass allein für die englische Fassung. (Issy van der Velde, 2025, GamesRadar)

Während ein Film nach maximal 3 Stunden vorbei ist, kann man in Videospielen gerne über 100 Stunden an Zeit reinstecken und zusätzlich ist das Medium interaktiv, was zu tausenden an verschiedenen Interaktionen führt. KI-Stimmen könnten hierbei die Möglichkeit bieten, die Welt noch lebendiger zu gestalten, da eine KI-Stimme wirklich auf alle Handlungen des Spielers reagieren kann, ähnlich wie sich ein unendliches Gespräch mit "ChatGPT" oder anderen KI-Chatbots führen lässt.

Die Gaming Community spielt bereits mit den modernen KI-Möglichkeiten herum, und kreierte so eine Modifikation für das 2011 erschienene Game "The Elder Scrolls V: Skyrim", durch welche man als Spieler mit den NPCs, welche durch die Welt laufen, frei reden konnte. Während der Spieler im Originalspiel nur eine limitierte Anzahl an Fragen und Antworten hatte, auf welche dann vorher festgelegte Reaktionen der NPCs kamen, machte es die Modifikation möglich, komplette Gespräche zu führen. In seinem YouTube Video "Wie Deutsch kann man Skyrim spielen?" sprach der Creator "RvNx Mango" so zum Beispiel mit einer Elfe über die Sicherheitsnormen der Brücke einer Fantasy Stadt. Ein solches Gespräch war für das Spiel eigentlich nie geplant, durch KI jetzt aber möglich gemacht. (@RvNxMango, 2026, YouTube)

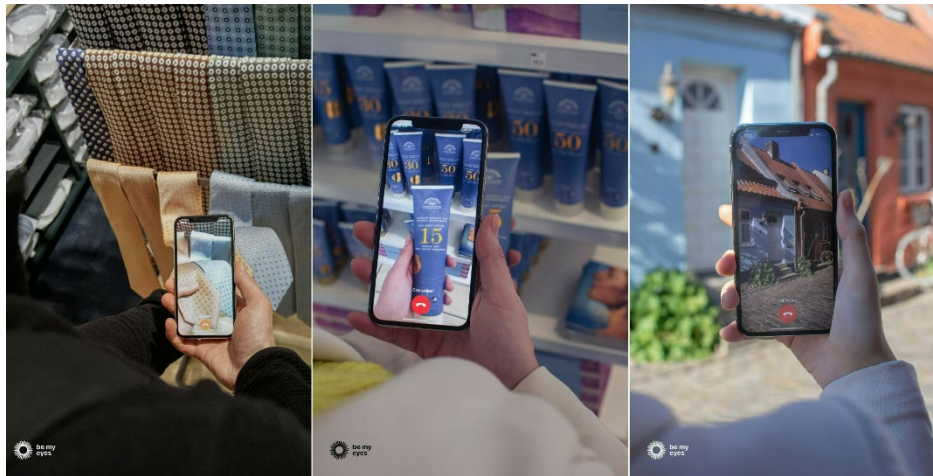
Abseits der Kunst gibt es aber auch praktische Anwendungen von neuen KI-Stimmen. Sprachassistenten wie Siri oder Alexa und auch die Sprachfunktionen der gängigen KI-Modelle wie ChatGPT oder Gemini sind mittlerweile im Alltag angekommen, und erleichtern die Kommunikation mit den KI-Systemen. ElevenLabs Gründer Mati Staniszewski glaubt sogar daran, dass Stimmen das Interface der Zukunft werden: *"what we think and believe from kind of the start is voice will fundamentally be the interface for interacting with technology"* (2026, Sequoia)

Visionen, wie das aussehen könnte, haben bereits diverse Filme über die Jahre gezeigt. Eine solche Kommunikation mit Technologie würde es nicht nur ermöglichen einfacher zu interagieren, ganz egal ob beim Spazieren, Fahrrad fahren oder sogar Auto fahren, es wäre auch wieder eine Möglichkeit für mehr Zugänglichkeit.



Superheld „Iron Man“ und sein sprachgesteuertes UI (Jon Favreau Iron Man, 2008, Paramount Pictures)

Be My Eyes ist eine App für Blinde und sehbehinderte Menschen, mit welcher es möglich ist, die Handykamera auf Texte, Gegenstände oder andere Sachen zu richten, und eine KI-Stimme beschreibt dann, was zu sehen ist. Eine unglaubliche Unterstützung, die nur durch KI-Stimmen überhaupt möglich ist. (be my eyes)



Anwendungsbeispiele für die App „Be My Eyes“ (Patient Innovation)

Mit Apple Voice Control oder Windows Voice Access ist zudem eine direkte Kommunikation mit Geräten und Apps möglich, was Menschen mit Querschnittslähmung oder fehlenden Gliedmaßen den Umgang mit Technologie erleichtert. (Ben Schwan, 2026, Heise)

Doch auch für gesunde Menschen bieten KI-Stimmen mehr Zugänglichkeit, zum Beispiel beim Kundensupport. Der Flughafen Berlin-Brandenburg setzt auf seinen KI-Stimmen-Agenten „Berry“ und erhielt dafür den Preis für die „Most Innovative Airport Initiative“ und erreichte eine Kundenzufriedenheit von 85%. (Mitglieder-News, 2025, CCV)

Die Firma Cognigy versorgt mit ihren KI-Agenten verschiedenste Unternehmen wie Toyota oder DHL. Oftmals gibt es dann immer noch die Möglichkeit einen menschlichen Support zu erhalten, vor allem wenn die KI nicht weiterweiß, aber die Möglichkeit von KI-Kundensupport übers Telefon kann einfache Fragen schnell klären und so die Arbeit erleichtern. (Cognigy)

Einer der größten Märkte für neue KI-Stimmen ist aber zweifelsohne KI-Musik. Laut einem Bericht von Billboard werden auf Suno, dem Marktführer im KI-Musik Bereich, täglich über 7 Millionen Songs generiert. (Kristin Robinson, 2025, Billboard)

Damit wird alle 2 Wochen ungefähr so viel Musik generiert, wie ganz Spotify umfasst. (Spotify) Mittlerweile erreicht diese Flut an KI-Songs auch den Streaminganbieter, der bereits angefangen hat über 75 Millionen dieser Songs zu löschen. (AI Landes, 2026, Gadget Review)

Auch, wenn es in diesem Abschnitt eigentlich nicht mehr ums Training gehen sollte, muss man hier doch kurz einen Artikel von The Atlantic erwähnen, der in einer riesigen Recherche tatsächlich zusammentragen konnte, an welcher Musik diese KI-Songs eigentlich trainiert wurden. (Alex Reisner, 2026, The Atlantic)

Bekannte Musiker wie LGoony sprachen sich nach dem Artikel auf den Sozialen Medien empört dagegen aus: „*Es basiert alles auf der Arbeit von anderen Leuten*“. Von LGoony lassen sich ganze 104 Songs in der Datenbank finden, manche davon sind nicht einmal auf den gängigen Streaminganbietern verfügbar. (@lgoony, 2026, Instagram)

Die Gema reichte auch eine Klage gegen Suno ein, die sie tatsächlich in einem jetzt als wegweisend geltenden Gerichtsurteil gewann, auch wenn das Urteil noch nicht final ist und Suno Einspruch erheben kann. *„Kein KI-Musikanbieter in Europa kann dieses Urteil ignorieren. Die Botschaft ist klar: Wer mit menschlicher Kreativität Geld verdient, muss die Menschen dahinter auch bezahlen“* (Zitat: Rechtsanwalt Christian Solmecke, 2026, Tagesschau)

Wie bei Social Media ist das Problem in der Nutzung die unglaubliche Schnelligkeit, in der sich Hunderte an Songs generieren lassen. Echte Künstler*Innen können unmöglich mit dieser Geschwindigkeit Schritt halten, kämpfen dann aber plötzlich in einem komplett überfluteten Markt um Aufmerksamkeit. Dabei ist dann komplett egal wie gut oder schlecht die KI-Musik tatsächlich ist, allein die Quantität hinter den Songs ist einfach ein Problem.

Eine wirkliche Begeisterung für diese KI-Songs ist auch nicht vorhanden. Zwar können in Tests mittlerweile 97% nicht mehr zwischen KI-Songs und echten unterscheiden, 80% fordern aber auch ein klares KI-Label, nur 11% halten es für richtig, dass KI-Musik genau gleich in den Charts behandelt wird, wie die von Menschen und über 73% halten das momentane Training der KI-Firmen für falsch. 66% würden KI-Musik zwar aus Neugier mal eine Chance geben und reinhören, 45% würden KI-Musik aber gerne komplett aus Streaming Anbietern herausfiltern. (Jesper Wendel, 2025, Deezer Newsroom)

Fazit

Die Nutzung neuer KI-Stimmen ergibt bisher die meisten Probleme. Die Schnelligkeit in der Social Media Content, KI-Musik oder Synchronfassungen erstellt werden können ist unschlagbar. Jetzt könnte man argumentieren, dass es ja egal ist, ob eine Musik oder Synchronfassung KI ist, wenn der Zuschauer keinen Unterschied erkennen kann. Diese Argumentation ist aber nicht ganz schlüssig, immerhin ist eine klare Ablehnung dieser Inhalte zu sehen, oftmals natürlich auch wegen der Qualität, aber auch wegen der Ethik dahinter. Zuschauer haben ein Recht darauf, zu erfahren, was KI ist und was nicht. Wenn man jetzt eine Schokolade im Supermarkt kauft, könnte eine Person auch nicht am Geschmack erkennen wie viel Zucker drin ist, ob diese durch Kinderarbeit entstanden ist und ob da eventuell tierische Produkte drin sind, die einen Veganer stören könnten. Genau dafür gibt es Regeln für die Produktbeschreibung, um den Konsumenten aufzuklären. Dasselbe Recht sollte auch für KI-Inhalte gelten, damit der Konsument selbst entscheiden kann.

Es gibt aber auf jeden Fall auch ganz neue, spannende Möglichkeiten, die sich mit den neuen KI-Stimmen ergeben. Moderne Interfaces und schnellerer Kundensupport sind nur 2 Beispiele, welche Arbeit auch nicht wegnehmen, sondern erleichtern, und in ihrer Existenz einen klaren Nutzen vorweisen. Wieder einmal bieten sich vor allem im Bereich Inklusion neue Möglichkeiten, um Sehbehinderten Menschen zu helfen und ihnen die Welt sowie diverse visuelle Medien zugänglicher zu machen.

Conclusio

Es scheint so, dass vor allem im Bereich der Kunst sind KI-Stimmen oftmals nur eine Traumvorstellung von großen Firmen sind, um sich Sprecher*innen zu sparen, ganz egal ob in Games, bei Musik, Synchronisationen, Hörbüchern und so weiter. Solche Umsetzungen stoßen momentan sowohl noch auf technische Hürden als auch auf eine fehlende Akzeptanz in der Bevölkerung, und die Gefahr eines Shitstorms ist groß. Beispiele wie Arc Raiders oder Deadly Patient zeigen das ganz gut.

Es ist auch wichtig festzuhalten, dass diese Stimmen nicht einfach so aus dem Nichts gekommen sind. Sie wurden trainiert, an Milliarden von Daten, die illegal verwendet wurden. Dass dann exakt diese Modelle genutzt werden sollen, um Menschen zu ersetzen, ohne die diese Modelle niemals möglich gewesen wären, und ohne, dass diese Menschen irgendwie davon profitieren, ist sehr kritisch zu betrachten.

Konrad Bösherz hat das in der Bachelor Doku nochmal aus seiner Perspektive ausformuliert: „Und das ist vielleicht auch der Unterschied dazu wenn es früher Kutscher gab und dann kamen die Autos dann gab es natürlich eine externe äußere Erfindung auf die du keinen Einfluss hattest [...] aber es hat niemand deine Arbeit als Kutscher sozusagen verwendet um deinen eigenen Job einfach zu tilgen“ (Florian Ploner & Timm Wehlte, 2026, Bachelor Doku)

KI-Stimmen sollten wohl keine echten Stimmen ersetzen, es raubt der Kunst die Kraft und nimmt ihr ihre Facetten und Nuancen und vor allem das Menschliche. „Dass KI einfach Synchronstimmen ist für mich irgendwie moralisch verwerflich auch ein Stück weit“ sagt so auch ein Besucher der Dokomi (Florian Ploner & Timm Wehlte, 2026, Bachelor Doku)

Es ist aber auch zu sehen, dass in der Kunst ganz neue Möglichkeiten entstehen, die Sprecher*innen nicht ersetzen, sondern mit ihnen zusammenarbeitet. So können spannende und immersive Erfahrungen erschaffen werden, die zuvor undenkbar waren, vor allem ein so komplexes und interaktives Medium wie Videospiele könnte so ganz neue Immersionen ermöglichen. Erste Spiele und Modifikationen zeigen, wo die Reise hingehen könnte, und dass Spieler durchaus gewillt sind, sich darauf einzulassen, wenn es die Kunst fördert und nicht nur kostensparender macht.

Es ist unter dem Aspekt auch möglich, KI-Stimmen auf ethische Weise zu trainieren, Deals auszuhandeln und Lizenzen zu kaufen, um ein faires Arbeitsumfeld für die Sprecher*Innen zu schaffen.

Doch vor allem abseits von der Kunst bilden sich unglaubliche Chancen in der Technologie. Die Möglichkeiten, die vielen Menschen mit Behinderungen dadurch gegeben werden, sind nicht nur faszinierend, sondern auch ein wichtiger Schritt und eine wahre Innovation. Vor allem, dass verstumme Menschen mit KI ihre eigene Stimme wieder bekommen können, ist unglaublich wichtig.

Außerdem können KI-Stimmen Prozesse optimieren und beschleunigen, vom Kundenservice bis hin zu modernen Interfaces, die ganz mit der Stimme kontrollierbar und interaktiv sind.

Problematisch sind jedoch die Betrugsmöglichkeiten und auch die massive Flut an „KI Slop“ die jetzt schon das Internet überrollt ist ein Problem, sowohl für Konsumenten als auch für die Kreativen. Es braucht klare Regeln, auch in der Online-Welt, und vor allem gegen

Deepfakes oder Identitätsdiebstahl muss vorgegangen werden. Hier ist es aber auch gut möglich, dass mit der Zeit immer mehr Aufklärung herrschen wird und die Leute ein Gefühl für diese Fakes entwickeln werden. Gesetze wie die jetzt entstandene Kennzeichnungspflicht von KI-Stimmen auf EU-Ebene sind auch ein wichtiger Schritt dabei. (Patrick Lipke, 2026, ZDF Heute)

Insgesamt lässt sich also sagen, dass KI-Stimmen ein komplexes Thema sind und auch noch länger bleiben werden. Die Chancen sind riesig, und viele davon sind nicht wirklich ethisch bedenklich, sondern bieten neue Wege die für die Menschheit vor ein paar Jahrzehnten noch utopisch gewesen wäre.

Genauso gibt es aber auch viele unethische Möglichkeiten, die durch KI-Stimmen entstehen. Ganz besonders im Vordergrund stehen das Training, welches geheim im Hintergrund stattfindet, ohne Transparenz, Ehrlichkeit oder Vergütung. Menschen sollten nicht ersetzt werden, und nur weil KI-Firmen jetzt stark für ihre Produkte propagieren, heißt es nicht, dass das auch die Zukunft sein muss, auf die sich die Menschen einlassen müssen. Erste Studien und Shitstorms nach KI-Nutzung zeigen das, ob sich das aber auch halten wird, ist eine Frage, die sich erst in den kommenden Jahren zeigen wird.

Literaturverzeichnis

- Q1 **BSI – Bundesamt für Sicherheit in der Informationstechnik** (2025): *Generative KI-Modelle* <https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/KI/Generative%20KI-Modelle.html>
- Q2 **Karen Hao** (2025): *Empire of AI*, S. 125, 131
- Q3 **ElevenLabs**: *Voice Cloning* <https://elevenlabs.io/de/voice-cloning>
- Q4 **Florian Ploner & Timm Wehlte** (2026): *Bachelor Doku* (Interviews mit Sebastian Deubelli, Anna-Sophia Lumpe, Marios Gavrilis, Sven Plate, Peter Flechtner & Konrad Bösherz)
- Q5 **Manfred Bremmer** (2026): *Was ist Scraping?*, Computerwoche <https://www.computerwoche.de/article/2804141/was-ist-scraping.html>
- Q6 **Abirami Vina** (2025): *A quick guide for beginners on how to train an AI model*, Ultralytics <https://www.ultralytics.com/de/blog/a-quick-guide-for-beginners-on-how-to-train-an-ai-model>
- Q7 **NVIDIA**: *What Is Generative AI?* <https://www.nvidia.com/en-us/glossary/generative-ai/>
- Q8 **ElevenLabs**: *Voice Cloning* <https://elevenlabs.io/voice-cloning>
- Q9 **Dustin Blank** (2025): *Announcing partnership with Sir Michael Caine*, ElevenLabs Blog <https://elevenlabs.io/blog/announcing-partnership-with-sir-michael-caine-to-newly-launched-iconic-marketplace>
- Q10 **Dustin Blank** (2025): *The Wonderful Wizard of Oz*, ElevenLabs Blog <https://elevenlabs.io/de/blog/the-wonderful-wizard-of-oz>
- Q11 **ElevenLabs**: *Voice Data Partnerships Application* <https://elevenlabs.io/de/voice-data-partnerships-application>
- Q12 **Julia Hercka** (2026): *Synchronsprecher als KI-Futter: Kampf um Stimmrechte*, NDR <https://www.ndr.de/kultur/film/synchronsprecher-als-ki-futter-kampf-um-stimmrechte,synchronsprecherki-100.html>
- Q13 **Kali Hays** (2023): *Marc Andreessen / Horowitz AI copyright*, Business Insider <https://www.businessinsider.com/marc-andreessen-horowitz-ai-copyright-2023-11>
- Q14 **Focus & Chip** (2025): *Telefonbetrug mit KI* https://www.focus.de/wissen/aufgepasst-telefonbetrug-mit-ki-so-entlarven-sie-gefaelschte-stimmen-von-familie-und-freunden_f536660e-38ca-4a16-b682-26302d216556.html
- Q15 **2 Bored Guys** (2026): *Der Absturz von Elon Musk* <https://www.youtube.com/watch?v=a-Kr1VYw0&t=1s>
- Q16 **@schlantologie** (2026): *YouTube-Kanal / KI-Satire: KI-Friedrich Merz als böser Imperator* <https://www.youtube.com/@schlantologie/videos> | <https://www.youtube.com/watch?v=xwFjWwBRQmI&t=20s>
- Q17 **ZDF Satire / Die Anstalt** (2026): *Friedrich Merz in: „Die böse Krankschreibung“ | Die Anstalt* <https://www.youtube.com/watch?v=IZAbfUqUQKs&t=2s>
- Q18 **Duden**: *Definition Parodie* <https://www.duden.de/rechtschreibung/Parodie>
- Q19 **Roman Schmidt & Fabian Reinholz** (2025): *Stimmklone durch KI vor Gericht: Synchronsprecher gewinnt vor dem LG Berlin*, Härting Rechtsanwälte <https://haerting.de/wissen/stimmklone-durch-ki-vor-gericht-synchronsprecher-gewinnt-vor-dem-lg-berlin/>
- Q20 **CloneMe**: *Case Study: Bringing History to Life with AI – The Democracy Matters Initiative* <https://www.cloneme.io/post/case-study-bringing-history-to-life-with-ai-the-democracy-matters-initiative>
- Q21 **Steam** (2025): *ARC Raiders Store Page* https://store.steampowered.com/app/1808500/ARC_Raiders/
- Q22 **Ari Notis** (2025): *Arc Raiders' use of AI voices is The Finals controversy all over again*, Polygon <https://www.polygon.com/arc-raiders-ai-voices-the-finals-embark-studios/>
- Q23 **Jamie Hore** (2025): *Arc Raiders dev "in no way uses generative AI whatsoever" for its visuals, but here's why Embark still has to disclose on Steam*, PCGamesN <https://www.pcgamesn.com/arc-raiders/generative-ai-use>
- Q24 **Zack Zwiezen** (2026): *Arc Raiders Has Started Replacing AI Voices With Human Ones Which CEO Admits Are Better*, Kotaku <https://kotaku.com/arc-raiders-replaced-ai-generated-content-human-recorded-dialogue-voices-2000678774>
- Q25 **ElevenLabs** (2025): *Lumiere Ventures and ElevenLabs collaborate to honor Alain Dorval in Sylvester Stallone's new film* <https://elevenlabs.io/blog/lumiere-ventures-and-elevenlabs-collaborate-on-sylvester-stallones-new-film>
- Q26 **Aurore Bergé** (2025): *X-Post / Tweet bezüglich der KI-Stimme von Alain Dorval* <https://x.com/auroreberge/status/1877847594254413885?s=20>
- Q27 **Eduard Altmann** (2026): *Telefonbetrug: 626.555 Fälle im Juni – KI-Betrug auf Rekordhoch*, Börse Express <https://www.boerse-express.com/news/articles/telefonbetrug-626555-faelle-im-juni-ki-betrug-auf-rekordhoch-928280>
- Q28 **@iBlali** (2026): *Die 10 verstörendsten Inhalte, die aus Spielen entfernt wurden | Videospielmythen* https://www.youtube.com/watch?v=Emwdz3N_CbY
- Q29 **@MrBeast** (2025): *World's Strongest Man Vs Robot* https://www.youtube.com/watch?v=Ah_uuTwGOYU
- Q30 **Bridging Voice / Scott-Morgan Foundation**: *1 Million Voices Initiative Partnership* <https://bridgingvoice.org/> | <https://bridgingvoice.org/elevenlabs-bridging-voice-partnership/>
- Q31 **Timothy Beck Werth** (2026): *ElevenLabs restored Eric Dane's voice with AI. Now they're offering 1 million voices for free*, Mashable <https://mashable.com/article/elevenlabs-eric-dane-1-million-voices-initiative-sxsw-2026>
- Q32 **ElevenLabs** (2024): *Our Impact Program aims to empower 1 million people through AI voice technology* <https://elevenlabs.io/blog/impact-program-announcement>
- Q33 **@TimGreen-NothingLeftUnsaid** (2026): *Podcast: Nothing Left Unsaid* <https://www.youtube.com/watch?v=iEBuNK66CGE&t=334s>
- Q34 **Andreea Bensa-Cruz** (2026): *Forscher klonen natürliche Stimmen mit Künstlicher Intelligenz*, Futurezone <https://futurezone.at/science/elektrolarynx-ki-stimmumwandlung-natuerliche-stimme-kehlkopfkrebs/403141736>
- Q35 **Lingraphica**: *What is an AAC Device?* <https://lingraphica.com/aac-devices/what-is-an-aac-device/>
- Q36 **Acapela Group**: *My-Own-Voice* <https://mov.acapela-group.com/de/mehr-erfahren/>
- Q37 **Caryn.ai** (2023): *Official Website* <https://www.caryn.ai/>

- Q38 **Alexandra Sternlicht** (2023): A 23-year-old Snapchat influencer used OpenAI's technology to create an A.I. version of herself that will be your girlfriend for \$1 per minute, Fortune <https://fortune.com/2023/05/09/snapchat-influencer-launches-carynai-virtual-girlfriend-bot-openai-gpt4/>
- Q39 **Joshua Zitser** (2023): Influencerin startet KI-Version von sich selbst: Abonnenten zahlen einen Dollar pro Minute, um mit ihr zu chatten, Business Insider <https://www.businessinsider.de/wirtschaft/international-business/influencerin-startet-ki-version-von-sich-selbst-abonnenten-zahlen-einen-dollar-pro-minute-um-mit-ih-er-zu-chatten/>
- Q40 **Ciriello, Raffaele and Hannon, Oliver and Chen, Angelina and Vaast, Emmanuelle**, January 05, 2024, *Ethical Tensions in Human-AI Companionship: A Dialectical Inquiry into Replika* Hawaii International Conference on System Sciences (HICSS 2024), Available at SSRN: <https://ssrn.com/abstract=5285198> or <http://dx.doi.org/10.2139/ssrn.5285198>
- Q41 **Der Spiegel** (2026): Google und Character.AI legen Klage wegen Suizid eines Teenagers bei https://www.spiegel.de/netzwelt/apps/google-und-character-ai-legen-klage-wegen-suizid-eines-teenagers-bei-a-e9e7c379-46d1-4c19-9e1e-55f957b028e2?sara_ref=re-xx-cp-sh
- Q42 **CBS News** (2025): Open AI, Microsoft sued over ChatGPT's alleged role in fueling man's "paranoid delusions" before murder-suicide in Connecticut <https://www.cbsnews.com/news/open-ai-microsoft-sued-chatgpt-murder-suicide-connecticut/>
- Q43 **Thibault Spirlet** (2026): KI-Chatbots sollen Einsamkeit heilen – Psychologieprofessor warnt vor dieser gefährlichen Nebenwirkung, Business Insider <https://www.businessinsider.de/leben/ki-gegen-einsamkeit-yale-professor-warnt-vor-dem-preis/>
- Q44 **Curtis, Liam** (2026): The TikTok AI Slop Report, Kapwing <https://www.kapwing.com/resources/the-tiktok-ai-slop-report/>
- Q45 **Hamilton, Eric** (2026): "AI slop" hurts consumers and creators. But high-quality AI could help both, University of Florida <https://news.ufl.edu/2026/03/ai-slop/>
- Q46 **@dieshowmitchris** (2026): Instagram-Reel zu KI-Synchro auf Prime Video <https://www.instagram.com/reel/Daa3mkqokGs/>
- Q47 **Uwe Mantel** (2026): Ich sollte da gestorben haben: Amazon löscht peinliche KI-Synchro, DWDL <https://www.dwdl.de/nachrichten/107081/ich-sollte-da-gestorben-haben-amazon-loescht-peinliche-kisynchro>
- Q48 **Anke Nicolai** (2020): Hören, was andere sehen – Audiodeskription: Zugang für blinde und sehbeeinträchtigte Menschen zu visuellen Medien, Netzwerk Kultur und Inklusion https://kultur-und-inklusion.net/wp-content/uploads/2020/01/Nicolai_Audiodeskription_2018.pdf
- Q49 **Agnieszka Szarkowska** (2011): Text-to-speech audio description. Towards wider availability of AD, The Journal of Specialised Translation, S. 156 <https://www.jostrans.org/article/view/7448>
- Q50 **Jonas Herrmann** (2026): Tomodachi Life: Wo Träume wahr werden im Test, GamePro <https://www.gamepro.de/artikel/tomodachi-life-wo-traeume-wahr-werden-im-test.3451376.html>
- Q51 **Issy Van der Velde** (2025): "I was in that sound booth for over 500 hours": Kingdom Come Deliverance 2 actor says the RPG's 2.2 million word script made recording an absolute saga, GamesRadar <https://www.gamesradar.com/games/rpg/i-was-in-that-sound-booth-for-over-500-hours-kingdom-come-deliverance-2-actor-says-the-rpgs-2-2-million-word-script-made-recording-an-absolute-saga/>
- Q52 **@RvNxMango** (2026): YouTube-Video: Wie Deutsch kann man Skyrim spielen? <https://www.youtube.com/watch?v=7zpjy4ft3Gw&t=419s>
- Q53 **Sequoia** (2026): Training Data: Mati Staniszewski, Podcast <https://sequoiacap.com/podcast/training-data-mati-staniszewski/>
- Q54 **Jon Favreau** (2008): Iron Man, Paramount Pictures
- Q55 **Be My Eyes**: Official Website & Patient Innovation Case <https://www.bemyeyes.com/> | <https://patient-innovation.com/post/1133?language=de>
- Q56 **Ben Schwan** (2026): Mit KI: Apple stellt neue Funktionen für Barrierefreiheit vor, Heise <https://www.heise.de/news/Mit-KI-Apple-stellt-neue-Funktionen-fuer-Barrierefreiheit-vor-11298952.html>
- Q57 **CCV / Mitglieder-News** (2025): Sechs Monate als AI Unicorn: Parloa knackt Umsatz-Marke von 50 Millionen US-Dollar <https://cc-verband.de/sechs-monate-als-ai-unicorn-parloa-knackt-umsatz-marke-von-50-millionen-us-dollar>
- Q58 **Cognigy**: Case Study DHL <https://www.cognigy.com/de/case-study/dhl>
- Q59 **Kristin Robinson** (2025): Suno Creates an Entire Spotify Catalog's Worth of Music Every Two Weeks, Says Investor Pitch Deck for \$250M Fundraise, Billboard <https://www.billboard.com/pro/suno-creates-spotify-catalog-music-two-weeks-pitch-deck/>
- Q60 **Spotify**: Company Info / Newsroom <https://newsroom.spotify.com/company-info>
- Q61 **AI Landes** (2026): Spotify Deleted 75 Million AI-Generated Tracks – and It's Not Done Yet, Gadget Review <https://www.gadgetreview.com/spotify-deleted-75-million-ai-generated-tracks-and-its-not-done-yet>
- Q62 **Alex Reisner** (2026): The Millions of Songs Mashed Into AI-Generated Music, The Atlantic <https://www.theatlantic.com/technology/2026/06/ai-music-generators-suno-google-udio/687485/>
- Q63 **@Igoony** (2026): Instagram-Reel zur Nutzung von Musikdaten für KI-Training <https://www.instagram.com/reels/DaLgDfBoAq0/>
- Q64 **Tagesschau / Solmecke, Christian** (2026): KI-Firma darf Melodien nicht ungefragt nutzen <https://www.tagesschau.de/kultur/urteil-gema-vs-suno-100.html>
- Q65 **Jesper Wendel** (2025) Deezer/Ipsos survey: 97% of people can't tell the difference between fully AI-generated and human made music – clear desire for transparency and fairness for artists, Deezer Newsroom <https://newsroom-deezer.com/2025/11/deezer-ipsos-survey-ai-music/>
- Q66 **Patrick Lipke** (2026): Wann Sie KI in Fotos, Videos oder Texten kennzeichnen müssen, ZDF Heute <https://www.zdfheute.de/ratgeber/ki-kuenstliche-intelligenz-ai-act-eu-100.html>

Alle Quellen wurden zuletzt am 03.08.2026 von 13-14 Uhr aufgerufen