

Quelle: <https://cloud.army/resources/blog/perception-of-ai-voices>

Übersetzt mit DeepL.com

WAHRNEHMUNG VON KI-STIMMEN

VON: DARREN BRIDGER, VIZEPRÄSIDENT, FORSCHUNG UND WISSENSCHAFT

Veröffentlicht am: 5. Juli 2023

Künstliche Stimmen, die mithilfe von KI erzeugt werden, sind auf dem besten Weg, ein wichtiges Instrument für Marken zu werden. Sie werden bereits in Telefon-Hotlines und interaktiven sprachbasierten Apps (wie Apples Siri und Amazons Alexa) eingesetzt und kommen zunehmend auch in der Werbung zum Einsatz. Doch wie wird sich der Einsatz solcher Stimmen auf die Wahrnehmung einer Marke durch die Verbraucher auswirken?

Das Konzept von sprechenden Computern und Robotern ist seit langem ein fester Bestandteil der Science-Fiction und zeigt eine Zukunft, in der Computer so flüssig kommunizieren wie Menschen. Als Pionier dieser Idee stellte Philip K. Dick in seinem Roman „Do Androids Dream of Electric Sheep?“ (auf dem der Film „Blade Runner“ basiert) Androiden, die von Menschen nicht zu unterscheiden waren, einschließlich ihrer Sprache. Filme wie „2001: Odyssee im Weltraum“ stellten den erschreckend eloquenten HAL 9000 vor, und „Star Wars“ schenkte uns C3PO, einen humanoiden KI mit menschenähnlicher Intelligenz und der Beherrschung Tausender Sprachen. Und im Fernsehen erweckte „Star Trek“ Data zum Leben, einen Androiden mit beeindruckender Fähigkeit zur menschlichen Sprache, und „Westworld“ schuf Roboter-Hosts, die menschliches Verhalten nahtlos simulieren.

In vielen dieser Beispiele sind die KI-Stimmen zwar flüssig, aber oft übermäßig formell, gestelzt oder ohne emotionale Nuancen. Diese Beispiele haben oft unsere Erwartungen an KI geprägt, aktuelle technologische Fortschritte vorweggenommen und inspirieren weiterhin unser Verständnis davon, was KI werden könnte. Für manche mag es daher überraschend sein, dass es bereits KI-Stimmen gibt, die vollkommen menschlich klingen und oft nicht von echten Stimmen zu unterscheiden sind, einschließlich emotionaler Nuancen. Im Gegensatz dazu befinden sich beispielsweise die Gesichter realistischer Roboter immer noch im „Uncanny Valley“ und wirken mit ihrem sehr realistischen Aussehen, das jedoch nicht ganz realistisch genug ist, eher beängstigend. Eine KI-Stimmenaufnahme kann heute jedoch wirklich als menschliche Stimme durchgehen.

KI-STIMMEN UND MARKEN

Stimmen werden für Marken immer wichtiger. Die Auswahl der richtigen Stimme ist für eine Marke natürlich von entscheidender Bedeutung. KI-Stimmen können besonders nützlich sein, wenn eine Marke Inhalte in mehreren Sprachen oder Dialekten produzieren muss, da KI Stimmen mit einer Vielzahl von Akzenten und Sprachmustern generieren kann. Außerdem kann KI eine einzige Stimme verwenden, die für die Persönlichkeit der Marke optimiert ist und dennoch in vielen Sprachen nahtlos fließend spricht. Der Einsatz von KI-Stimmen ermöglicht zudem eine größere Flexibilität und Skalierbarkeit, da KI-Stimmen nach Bedarf generiert und modifiziert werden können.

Marken, die KI-Stimmen einsetzen, profitieren über Skalierbarkeit und Vielseitigkeit hinaus in mehrfacher Hinsicht. KI-Stimmen sind oft kostengünstiger als die Beauftragung professioneller Sprecher, insbesondere bei großen Projekten oder solchen, die mehrere Sprachen erfordern. Das eröffnet kleineren Marken mit bescheideneren Budgets die Möglichkeit, in Online-Anzeigen professionell klingende Audioinhalte zu

verwenden. Der Einsatz von KI gewährleistet außerdem eine konsistente Sprachqualität und einen einheitlichen Tonfall über verschiedene Inhalte hinweg und stärkt so die Markenidentität, wo menschliche Stimmen eher Schwankungen unterliegen.

KI-Stimmen können akribisch angepasst werden, um bestimmte Klänge oder Töne zu erzielen, die der Identität der Marke oder dem Kontext der Botschaft entsprechen – etwas, das mit menschlichen Stimmen schwierig sein kann. Darüber hinaus signalisiert der Einsatz von KI-Stimmen die innovative und technologisch fortschrittliche Haltung einer Marke, was potenziell ein bestimmtes Zielpublikum ansprechen kann. Manchmal kann es für eine Marke sogar von Vorteil sein, eine Stimme deutlicher als KI erkennbar zu machen, einfach um einen Kontrast zu ihrer ansonsten makellosen Simulation der menschlichen Stimme zu schaffen.

Die Verwendung von KI-Stimmen birgt jedoch potenzielle Risiken, die in Verbindung mit Transparenz das Image einer Marke beeinträchtigen können. Erstens könnten Kunden Marken, die KI-Stimmen verwenden, als unpersönlich oder emotionslos empfinden, was zu einem Verlust an Menschlichkeit führen kann – ein Faktor, der in Branchen oder Situationen, in denen zwischenmenschliche Beziehungen und Empathie eine wichtige Rolle spielen, von entscheidender Bedeutung ist. Zweitens könnten ethische Fragen aufkommen, da einige Kunden Bedenken äußern, dass KI möglicherweise Arbeitsplätze für Synchronsprecher verdrängen könnte.

Schließlich kann auch eine falsche Darstellung zum Problem werden. Wenn eine Marke KI-Stimmen ohne klare Angabe verwendet und die Verbraucher dies später entdecken, könnte dies zu einem Gefühl der Täuschung führen und den Ruf der Marke schädigen.

IMPLIZITE WAHRNEHMUNGEN VON KI-STIMMEN

Stimmen spielen eine wichtige Rolle bei unserem ersten Eindruck von einer Person. Eine Stimme kann Ihnen etwas über die körperlichen Merkmale einer Person verraten, wie beispielsweise ihre Größe und ihr Alter. Und natürlich auch über ihre psychologischen Eigenschaften. Wir sind sehr geschickt darin, uns allein anhand der Stimme eines Menschen innerhalb von 400 Millisekunden⁽¹⁾ einen ersten Eindruck zu verschaffen. Die beiden Eigenschaften, die bei der Bildung eines ersten Eindrucks von einer Stimme am wichtigsten zu sein scheinen, sind Vertrauen und Dominanz⁽²⁾.

Aus evolutionärer Sicht waren diese Eigenschaften wichtig, da sie unseren Vorfahren halfen, Fremde schnell einzuschätzen und zu entscheiden, ob sie eine Bedrohung darstellten. Dieser Prozess abläuft so schnell und größtenteils außerhalb unseres Bewusstseins, dass es uns fast unmöglich ist, ihn verbal zu beschreiben. Das macht die Wahrnehmung von Stimmen zu einem idealen Thema für implizite Forschung.

Bei CloudArmy haben wir eine Reihe von impliziten Tests mit zeitlich begrenzten Antworten durchgeführt, um die Wahrnehmung von KI-Stimmen in der Werbung zu untersuchen. Da KI-Stimmen mittlerweile kaum noch von echten Stimmen zu unterscheiden sind, lautet die eigentliche Frage für die Werbung unserer Meinung nach: Sollten Marken offenlegen, dass eine Stimme von einer KI generiert wurde? Sowohl die Offenlegung als auch die Verschleierung bergen Risiken.

Unsere Studie, die in Zusammenarbeit mit dem Audioexperten Steve Keller durchgeführt wurde und auf dem diesjährigen Nudgestock Festival vorgestellt wird, verwendete eine Reihe von Audioanzeigen für erfundene Marken (damit die teilnehmenden Verbraucher keine Vorurteile gegenüber den Marken hatten), die sowohl von menschlichen Sprechern als auch von KI-Stimmen aufgenommen wurden. Wir variierten systematisch, ob den Teilnehmern mitgeteilt wurde, dass sie eine KI-Stimme hören würden.

Wir haben festgestellt, dass zwar eine knappe Mehrheit (60 %) der Menschen glaubt, unterscheiden zu können, ob eine Stimme von einer KI oder einem Menschen stammt, aber als wir ihre Fähigkeit dazu überprüft haben, waren sie nicht besser als der Zufall!

Obwohl sie nicht bewusst erkennen konnten, ob eine Stimme von einer KI stammte, zeigten ihre impliziten Reaktionen ein anderes Bild.

Ihre impliziten Reaktionen auf eine Marke waren, nachdem sie eine von einer KI-Stimme vorgelesene Werbung gehört hatten, deutlich weniger positiv und vertrauensvoll als die Reaktionen derjenigen, die dieselbe Werbung von einer menschlichen Stimme vorgelesen bekommen hatten. (Dabei ist zu beachten, dass die KI-Stimme bewusst nicht als solche erkennbar war.)

Darüber hinaus führte bereits die Information, dass die Werbung, die sie hören würden, von einer KI gesprochen wurde, zu einer weniger positiven impliziten Reaktion auf die Marke – unabhängig davon, ob die Werbung von einem Menschen oder einer KI gesprochen wurde. Das Wissen um die KI prägte die implizite Reaktion.

Mit anderen Worten: Obwohl Menschen eine KI-Stimme nicht zuverlässig von einer menschlichen Stimme unterscheiden können, reagieren sie unbewusst negativer darauf. Und wenn ihnen gesagt wird, dass es sich um eine KI handelt, bleibt diese negative Reaktion bestehen, selbst wenn die Stimme menschlich war.

Um unsere Nudgestock 2023-Präsentation auf YouTube anzusehen, klicken Sie [hier](#).

Die Welt der KI-Stimmen ist ein sich ständig weiterentwickelndes Feld, das voller Möglichkeiten steckt. Für zukunftsorientierte Marken können diese Tools spannende neue Möglichkeiten bieten, um mit ihrem Publikum in Kontakt zu treten und ihre Markenidentität zu vermitteln. Für eine erfolgreiche Integration ist es jedoch entscheidend, die Wahrnehmung von KI-Stimmen durch die Verbraucher zu verstehen und zu steuern. Während wir weiterhin die Grenzen der KI-Technologie ausloten, werden Untersuchungen wie die von CloudArmy entscheidend dazu beitragen, ihre Rolle in der Markenkommunikation zu definieren.

References

(1) Mileva, M. and Lavan, N., 2023. Trait impressions from voices are formed rapidly within 400 ms of exposure. */Journal of Experimental Psychology: General/.*

(2) McAleer, P., Todorov, A. and Belin, P., 2014. How do you say 'Hello'? Personality impressions from brief novel voices. */PloS one/, /9/(3), p.e90779.*

Referenzen

(1) Mileva, M. und Lavan, N., 2023. Eindrücke von Charaktereigenschaften anhand von Stimmen entstehen innerhalb von 400 ms nach dem Hören. */Journal of Experimental Psychology: General/.*

(2) McAleer, P., Todorov, A. und Belin, P., 2014. Wie sagen Sie „Hallo“? Persönlichkeitsimpressionen aus kurzen, unbekanntem Stimmen. */PloS one/, /9/(3), S. e90779.*

