

Warum es bei Audio-KI in der Markenkommunikation auf die Strategie ankommt

von **Lars Ohlendorf, WeSound** Donnerstag, 02. November 2023



Inmitten des KI-Hypes stehen Markenverantwortliche vor der Herausforderung, die richtigen Entscheidungen für ihre Marke zu treffen. Zahlreiche Audio-KI Tools versprechen das akustische Markenerlebnis auf ein neues Niveau zu heben. Doch welche Tools liefern echten Mehrwert? Lars Ohlendorf, Head of Design bei WeSound, erklärt in seinem Gastbeitrag, warum Audio Branding und Audio Marketing weniger von den KI-Tools an sich, als vielmehr von einer strategischen Neuausrichtung profitieren.

Audio-KI eröffnet Unternehmen mit der Analyse und Erzeugung von Klängen nicht nur vielfältige Anwendungen und Perspektiven, sondern auch eine Reihe von Herausforderungen. Dennoch ist es schön zu sehen, wie viel Neues gerade ausprobiert und in kleinen Projekten entwickelt wird. Auch machen die öffentlichen Debatten Freude, zwischen "Wir werden alle arbeitslos (<https://www.horizont.net/medien/nachrichten/es-wird-krass-werden-wie-radikal-ki-die-audio-werbe-welt-veraendern-wird-211541>)" bis "Der größte Wachstumsmarkt seit Erfindung des Internets (<https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier>)". Hier bewegt sich viel, und es bewegt sich schnell.

Es scheint mir aber, dass reale Einsatzbeispiele mit Audio-KI eher selten diskutiert werden. Was nutzt zum Beispiel eine generierte Stimme, wenn sie für die Sendung in TV oder Radio (still a thing!) qualitativ nicht ausreicht? Gerade wenn wöchentlich neue Produkte und Services rund um Audio-KI in den Markt gedrückt werden, lohnt es sich, kurz von den Werkzeugen weg, hin zum Zweck ihres Einsatzes und damit zu einer möglichen Strategie zu schauen. Erst durch die

Kombination von realistischer Erwartungshaltung, technischem Verständnis und kreativen Strategien lassen sich Brücken von der Vorstellung zur praktischen Umsetzung schlagen und wirklich neue Kommunikationsmöglichkeiten und Erlebnisräume für Marken erschließen.

Das Versprechen: mit unzähligen Audio-KI Lösungen Klänge analysieren und erzeugen

Audio-KI entfaltet sich ungefähr entlang der beiden Hauptzweige Analyse und Erzeugung von Klängen. In Kommunikation und Werbung hilft Audio-KI dabei, komplexe Schalle wie Musiken, Stimmen und Sound Designs auf ihren akustischen Markenfit zu untersuchen, sowie kognitive und wahrnehmungsspezifische Faktoren wie Einprägbarkeit, Aktivierungsgrad und emotionale Valenz zu messen. Das stellt Entscheidern neben der Expertise der jeweiligen Design-Teams und dem eigenen Bauchgefühl eine günstige, wenngleich oft fehlerbehaftete, quasi-objektive zweite Meinung zur Verfügung – sagen wir, eine Marktforschung light – und eröffnet damit auch eine Reihe von ergänzenden Anwendungen wie musikalische Ähnlichkeitssuchen oder Abgleich mit musikalischen Trends in sozialen Medien. Gleichmaßen können Audio-KI-Systeme auch für die Analyse von komplexen Audio Echtzeitanwendungen genutzt werden: zum Beispiel für die Stimmenanalyse, insbesondere im Kontext von Call-Centern. Allein auf Basis der Stimme kann die KI Rückschlüsse auf die Gefühlslage des Anrufers ziehen und, in Verbindung mit generativer KI, markenkonforme Handlungsempfehlungen für den Call-Center-Agenten generieren.

Die *generative* Audio-KI hingegen ermöglicht die Produktion von Audio-Inhalten. Anwendungsfälle umfassen die komplette Generierung von Audioanzeigen mit Stimmen und Musik (vgl. Aflorithmic (<https://audiostack.ai/>), Adthos (<https://www.adthos.com/>), Omnicon's SARA (<https://www.horizont.net/tech/nachrichten/omg-projekt-sara-wie-omnicom-audiowerbung-mit-ki-zielgenauer-macht-205661>)). Sonic Branding Elemente können in Echtzeit generiert werden. In der AV Medien Produktion erweisen sich Voice Modelle als hilfreich und zeitsparend. Ebenso werden generierte Stimmen bereits für Podcasts oder Nachrichtenmeldungen verwendet. Dies ist nur ein kleiner Ausschnitt des aktuell Möglichen, und die Möglichkeiten werden im Wochentakt besser und mehr.

Wie hilfreich sind Audio-KI Tools wirklich?

Allerdings steht die Zahl der Möglichkeiten ihrer tatsächlichen Relevanz und Funktionalität für ein spezifisches Unternehmen diametral gegenüber: Hier müssen wir Audio-KI substanziellen Realitätsprüfungen unterziehen.

Beispielsweise erreichen generative Systeme, ob in der Spracherzeugung oder Musikproduktion, noch nicht die nötige technische und ästhetische Qualität, um neben echten Aufnahmen überzeugen zu können. Ebenso technischer Qualität ist die die korrekte Aussprache von Fachtermini, Wortschöpfungen oder kontextuell unterschiedlich betonter Worte. Je nach Algorithmus kann die Fehlerrate bei generierter Sprache durchaus bei 1 Prozent liegen. Wenn eins von hundert Worten potenziell falsch ausgesprochen wird, kann so ein System in automatisierten Szenarien noch nicht infrage kommen.

» Generative Systeme, ob in der Spracherzeugung oder Musikproduktion, erreichen noch nicht die nötige technische und ästhetische Qualität, um neben echten Aufnahmen überzeugen zu können. «

Lars Ohlendorf

Zudem wird man auf der inhaltlichen Seite sowohl von Sprach- als auch Musikmodellen noch lange nicht so etwas wie kreative Absicht hören: die künstliche Stimmelmelodie irrt ziellos, gleichermaßen folgen generierte Musiken keiner sinnvollen Bewegung. Als Zuhörer schaltet man schnell auf Durchzug – nicht unbedingt, was man sich als werbetreibendes Unternehmen wünscht. Zwar ließe sich hier durch aufwändiges Prompting gegensteuern, jedoch entstünde hierdurch so hoher Mehraufwand, dass sich jeder Gedanke an Einsparung oder Automation verlieren würde. Man würde lediglich Arbeit von einem Ort an einen anderen schieben.

Auf der anderen Seite weisen analytische Systeme, also maschinelles Lernen, bekanntermaßen eine ganze Reihe rechtlicher, wirtschaftlicher, strategischer und ethischer Fragen auf. Auf viele dieser Fragen warten wir noch auf Antworten aus der Politik, insbesondere, aber nicht nur in Deutschland (<https://www.nytimes.com/2023/05/03/opinion/ai-lina-khan-ftc-technology.html?searchResultPosition=4>). Solange diese Antworten aber fehlen, können und müssen Unternehmen diese selbst definieren; mit genauem Blick auf mögliche rechtliche Entwicklungen, die Außenwirkung, aber eben auch auf die wirtschaftlichen Ziele. Bei den tatsächlichen Anwendungen analytischer Audio-KI können wir die oben skizzierten Szenarien betrachten: Für die ad hoc Analyse von Musiken, Sprache und Sound Designs werden Unmengen an Daten genutzt, deren Herkunft oft nicht klar ist und deren Validität für einen bestimmten Markt nicht unbedingt gegeben ist. Für die Call Center Anwendung würde jede Kommunikation zwischen Agent*in und Kund*in analysiert werden. – Wie verhindert man hier Diskriminierung? Was macht der Agent, wenn der Algorithmus falsche Empfehlungen gibt? Darüber hinaus bedürfen, was angesichts der vielen „schlüsselartigen“ Lösungen der generativen Audio-KI leicht in den Hintergrund gerät, die Entwicklung und Implementierung dieser spezialisierten Systeme erhebliche Ressourcen.

Ausblick: Audio-KI braucht strategische Implementierung – schnelle Lösungen sind meist keine

Die Implementierung von Audio-KI wird also, abgesehen von wenigen spezifischen Einzelanwendungen, anfänglich keine Kosten einsparen, sondern bedeutet vielmehr wesentliche Investitionen. Dafür aber stehen völlig neue Interaktionsformen zwischen Marken und Kunden in Aussicht.

In der Fortsetzung der Auseinandersetzung mit Audio-KI gilt es, die Brücke zwischen der ideellen Vorstellung und der realen Anwendbarkeit zu schlagen. Wesentlich ist hier die Strategie, die wir in Workshops mit unseren Kunden methodisch entwickeln:

- Das Wunschscenario: Wie könnten analytische oder generative Audio KI-Verfahren tatsächlich Prozesse optimieren, Effizienz steigern und Effektivität? Gibt es alternative Verfahren?
- Die Ausgangsanalyse: Welche Daten braucht man dafür? Welche fertigen Tools helfen dabei, welche müssen entwickelt werden? Welche internen und externen Arbeitsgruppen werden benötigt?
- Die Projektion: Welche rechtlichen und ethischen Implikationen hat das? Wie werden Kund*innen und Mitarbeiter*innen reagieren? Wie optimiert man das System, auf welcher Grundlage optimiert es sich selbst? Wie befähigt man Arbeitnehmer*innen, mit den neuen Systemen umzugehen?

Die Antworten auf diese Fragen sind höchst individuell und machen immer wieder deutlich, dass es bei Audio-KI weniger um Tools an sich als vielmehr um neue Strategien auf Basis neuer, bisher nicht existenter Möglichkeiten geht. Mit einer ausgewogenen Kombination aus realistischer Erwartungshaltung, technischem Verständnis und kreativen Strategien können wir den Weg für eine sinnvolle Integration von Audio-KI in die Unternehmenskommunikation und Markenerlebnisräume ebnen. Dafür braucht man interdisziplinäre Teams: Es reicht nicht, wenn sich Branding Department und Design an einen Tisch setzen, sondern es müssen alle potenziell betroffenen Abteilungen und Personen einbezogen werden – nur so kann man dem Streetlight-Effekt entgegenwirken, nämlich nur dort nach Lösungen zu suchen, wo die thematische Lampe gerade hinleuchtet. Wenn dies aber gelingt, können ganz neue Kommunikationsformen entstehen, neue Formen der individuellen Ansprache und neue Erlebnisräume: Für neugierige und innovative Marken, die jetzt die richtigen Fragen stellen.