AIVINOVA

Audio Branding mit KI

Wie Unternehmen Voice AI nutzen können – Anwendungsfälle, Tools und Empfehlungen für modernes B2B-Marketing

Powered by:

digital quality
Next level to success

Executive Summary

Die Identität einer Marke umfasst weit mehr als nur ihre visuelle Gestaltung, auch der Klang der Marke prägt die Wahrnehmung nachhaltig. Voice AI bietet B2B-Unternehmen neue Möglichkeiten, ihre Markenidentität durch konsistente, skalierbare Audioinhalte zu stärken. Ob Social Media Posts, Produktvideos oder mehrsprachiger Content, synthetische Markenstimmen machen die Tonalität und den Charakter einer Marke als sogenannte Voice Avatare hörbar und wiedererkennbar. Vor dem Hintergrund eines zunehmend wachsenden Markts für KI-generierte Audio-Inhalte und der wachsenden Bedeutung auditiver Markenwahrnehmung ist es wichtig zu wissen, wie Sie Audio Branding gezielt in der Praxis einsetzen.

Dieses Whitepaper zeigt Ihnen, wie Sie Voice AI in verschiedenen Use Cases einsetzen können und welche Tools sich in der Praxis dazu bereits bewährt haben. Sie erhalten einen Praxisleitfaden zur Schritt-für-Schritt-Umsetzung einer KI-generierten Markenstimme. Voice AI ist damit für Sie nicht nur ein Effizienztreiber, sondern ein modernes Audio Branding Instrument im digitalen B2B-Marketing.

Warum der Klang der Marke für die Kundenbindung wichtig ist

Corporate Identity umfasst weit mehr als nur die visuelle Gestaltung einer Marke, auch ihr Klang prägt die Wahrnehmung bei Kunden nachhaltig. Genau hier setzt Audio Branding an: Die Entwicklung einer akustischen Markenidentität, die Wiedererkennung schafft und Emotionen auslöst. Mit dem Fortschritt von Voice-AI-Technologien, insbesondere Text-to-Speech-Systemen und KI-basierten Voice Avataren, erhalten Unternehmen heute die Möglichkeit, ihre Markenstimme konsistent, skalierbar und mehrsprachig über digitale Kanäle hinweg einzusetzen.

Zugleich entwickelt sich der Markt: KI-generierte Audio-Inhalte erzielten 2024 ein Volumen von 412 Mio. US-Dollar, bis 2033 wird ein Anstieg auf 3,43 Mrd. US-Dollar erwartet – eine jährliche Wachstumsrate von 25,2 %. Getrieben wird diese Entwicklung von der fortschreitenden Digitalisierung, der wachsenden Bedeutung auditiver Markenwahrnehmung und der hohen Skalierbarkeit KI-generierter Inhalte. Fortschrite in Generative AI und Deep Neuoral Network und die Möglichkeit, Audio-Inhalte in Echtzeit an Zielgruppenpräferenzen anzupassen, tragen zu dieser Entwicklung bei.

In der Praxis zeigt sich die Reife der Technologie bereits heute: Im Radio, in Podcaste oder auf Social Media kommt Voice AI etwa für KI-generierte Werbung oder die Umwandlung von Textbeiträgen in Audio-Inhalte zum Einsatz. Gleichzeitig stellt die Integration von Voice AI Unternehmen vor neue Herausforderungen: Die Auswahl geeigneter Tools, die Definition einer authentischen Markenstimme sowie die rechtssichere und strategiekonforme Einbettung in bestehende Marketingprozesse erfordern klare Leitlinien.

Wie Voice AI eingesetzt wird

Typische Einsatzmöglichkeiten für Voice AI sind:

Vertonung von Textposts:
Durch Text-to-Speech-Technologien lassen sich klassische Textposts ohne großen Aufwand in Audioformate umwandeln (z.B. als Kurzvideos auf LinkedIn).

Untertitel für Videos:
Da Social Media Videos heute häufig ohne Ton konsumiert werden, sind Untertitel essenziell. Voice AI generiert Untertitel schnell und einfach.

Voiceover für Erklär- und Produktvideos:
Die Videoproduktionen nutzt häufig noch professionelle Sprecher- Das ist mit Aufwand und Kosten verbunden. Durch Voice AI werden externe Sprecher redundant und die Produktionsdauer signifikant verkürzt.

Mehrsprachiger Content:
KI ermöglicht es, Audio-Inhalte schnell und einfach in verschiedenen Sprachen zu produzieren.

Markenstimmen mit Voice Avataren zum Leben erwecken

Ein Voice Avatar ist eine KI-basierte, digitale Markenstimme, die Sprache, Tonalität und Charakter einer Marke hörbar macht. Er kann in Form eines visuellen, sprechenden Avatars auftreten, also mit Mimik und Gestik. Ebenso zählt aber auch bereits die wiederholte Verwendung einer spezifischen, markentypischen synthetischen Stimme ohne sichtbare Darstellung als Voice Avatar. Entscheidend ist, dass eine konsistente Stimmpersönlichkeit geschaffen wird, die die Identität der Marke transportiert und Wiedererkennung schafft. Dabei lassen sich alle zuvor genannten Content-Formate, von vertonten Textposts über Erklärvideos bis hin zu mehrsprachigem oder barrierefreiem Content, mit Voice Avataren gut umsetzen.

Ein zentraler Baustein eines Voice Avatars ist die Entwicklung einer Markenstimme, das akustische Pendant zur visuellen Corporate Identity. Klang wirkt emotional und bleibt im Gedächtnis, da Musik und Geräusche in Gehirnregionen verarbeitet werden, die eng mit Erinnerung und Gefühl verknüpft sind. Ein klassisches Beispiel ist McDonald's die mit ihrem „I'm Lovin' It" eine globale Markensignatur aufgebaut haben. Gerade wenn Voice Content ein wichtiger Teil einer Social-Media-Strategie wird, ist es wichtig, dass auch der auditive Auftritt der Marke stimmig, professionell und wiedererkennbar bleibt. Ein Voice Avatar, abgestimmt auf die eigene Markentonalität, kann hier als zentrales Sprachrohr dienen.

Wie Sie das passende Voice AI Tool finden

Um die passende Voice-AI-Lösung zu finden, ist ein strukturierter Vergleich anhand funktionaler, technischer und wirtschaftlich-rechtlicher Kriterien entscheidend. Der folgende Kriterienkatalog bietet Ihnen eine fundierte Grundlage für die Bewertung und Auswahl geeigneter Tools.

Art	Kriterien	Beschreibung
Stimme und Text	Stimmqualität	Natürlichkeit der Stimme, Betonung, Sprachmelodie und Pausensetzung
	Stimmenvielfalt	Breite Auswahl an Stimmen in Bezug auf Geschlecht, Tonlage, Stil und Emotion
	Stimmenklon	Imitierung bestehender Stimmen
	Individualisierung	Einstellungen zur Individualisierung der Stimme (z.B. Stimmfarbe, Schnelligkeit, Akzente)
	Mehrsprachigkeit	Verfügbarkeit und Qualität verschiedener Sprachen
	Visuelle Synchronisation	Präzision der Lippenbewegung, Mimik und Timing bei visuellen Avataren
	Untertitel-Export	Automatisierte Erstellung von Untertiteln
	Übersetzung	Automatisierte Übersetzung von Texten
Benutzung und Integration	Bedien-freundlichkeit	Intuitive Benutzeroberfläche mit Auswahl an Templates und einfache Bedienung ohne Programmierkenntnisse
	Exportformate	Verfügbarkeit gängiger Dateiformate (z.B. MP4, WAV, SRT)
	Kompatibilität	Möglichkeit zur Integration in bestehende Tools oder automatisierte Workflows (z. B. via API).
Preis und Rechts-konformität	Preismodell	Art des Preismodells (z.B. Pay-per-Use, Staffelpreise)
	Datenschutz-konformität	Einhaltung rechtlicher Vorgaben, insbesondere bei EU-Nutzung (z. B. Serverstandorte, Datenverarbeitung)
	Kommerzielle Nutzungsrechte	Klare Regelung, ob und wie Inhalte kommerziell verwendet, bearbeitet und veröffentlicht werden dürfen

Das richtige Voice AI Tool für Sie

Die Auswahl an Voice AI Tools ist groß. Im Folgenden werden die vier bekanntesten Anbieter mithilfe der zuvor definierten Auswahlkriterien kurz gegenübergestellt. Diese Gegenüberstellung hilft Ihnen dabei das für Sie passende Tool auswählen zu können.

Kriterium	Amazon Polly	Murf AI	Lovo AI	Synthesia
Stimm-qualität	🟡	🟢	🟡	🟢
Stimmen-vielfalt	🔴	🟢	🟢	🟢
Stimmen-klon	-	🟢	🟢	🟢
Individua-lisierung	🟡	🟢	🟡	🟡
Mehrsprachig-keit	🟡	🟡	🟢	🟢
Visuelle Syn-chronisation	🔴	🔴	🟢	🟢
Untertitel-Export	-	🟢	🟢	🟢
Über-setzung	-	-	🟢	🟢
Bedien-freundlichkeit	🔴	🟢	🔴	🟢
Export-formate	🟢	🟢	🟡	🟡
Kompatibilität	🟢	🟢	-	🟢
Datenschutz-konformität	🟢	🟢	🟢	🟢
Kommerzielle Nutzungsrechte	🟢	🟢	🟢	🟢
Preismodell	Nutzungs-abhängiges Modell	Freemium mit gestaffeltem Abonnement	Freemium mit gestaffeltem Abonnement	Freemium mit gestaffeltem Abonnement

🟢 = Hohe Qualität/Auswahl 🟡 = Moderate Qualität/Auswahl 🔴 = Geringe Qualität/Auswahl

Ihre Markenstimme mit KI gezielt entwickeln

Audio Branding lebt von einer klar definierten Markenstimme, die in einem sogenannten Voice-AI-Styleguide definiert ist. Dieser sorgt dafür, dass Tonfall, Ausdruck und Format in allen vertonten Inhalten wiedererkennbar bleiben. Folgendes Prompt-Template kann bei der Erstellung eines Voice Avatars je nach verwendetem Voice AI Tool helfen.

1. Tonalität

Sie können den Tone of Voice anhand verschiedener Facetten beschreiben. Diese können Sie mit gezielten Anweisungen, nachfolgend als [*Prompt-Bestandteile*] dargestellt, im entsprechenden Voice AI Tool gezielt ansteuern, z.B. mit:

- **Formalität:** z.B. [*Sehr formell*] oder [*Locker-formell*] oder [*Zwanglos und umgangssprachlich*]
- **Seriosität:** z.B. [*Ernst und seriös*] oder [*Heiter und humorvoll*]
- **Respekt:** z.B. [*Respektvoll und einfühlsam*] oder [*Frech und unkonventionell*]
- **Motivation:** z.B. [*Nüchtern und sachlich*] oder [*Begeistert und leidenschaftlich*]
- **Stimmfarbe:** z.B. [*männlich und alt*] oder [*weiblich und jung*]

2. Sprachstil

Auch der Sprachstil kann mithilfe von Prompts individualisiert werden, beispielsweise durch:

- **Satzstruktur:** z.B. [*Sätze kurz und prägnant*] oder [*Sätze erklärend*]
- **Wortwahl:** z.B. [*Fachterminologie verwenden*] oder [*Fachbegriffe erklären*] oder [*Klare und direkte Wortwahl*]
- **Perspektive:** z.B. [*Ich-Form*] oder [*Sie-Form*] oder [*Neutrale Form*]
- **Fremdwörter:** z.B. [*Fremdwörter vermeiden*] oder [*Gezielter Einsatz von Fremdwörtern*] oder [*Fremdwörter einsetzen und erklären*]
- **Gendern:** z.B. [*Konsequent gendern*] oder [*Genderneutral formulieren*] oder [*Auf Gendern verzichten*]

3. Format

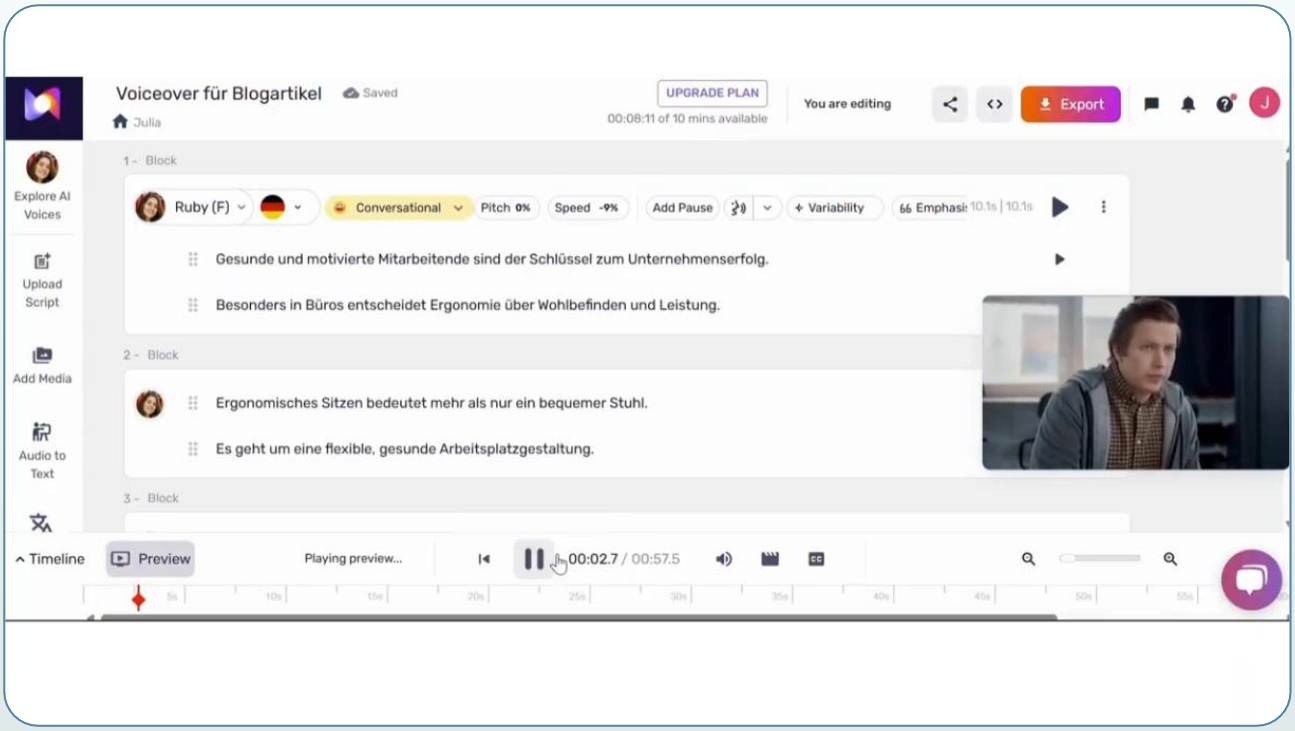
Weiter können Sie Charakteristika des Formats festlegen, die entweder generell oder für einzelne Marketingmaßnahmen gelten, wie z.B.:

- **Kanal:** z.B. [*LinkedIn Video*] oder [*Produktvideo für Webseite*]
- **Dauer:** z.B. [*30-90 Sekunden*] oder [*5 Minuten*]
- **Zielgruppe:** z.B. [*Fach- und Führungskräfte*] oder [*Pharmabranche*]

Kombinieren Sie die verschiedenen Prompt-Bestandteile miteinander, so erhalten Sie einen vollumfänglichen Prompt, mit dem Sie der Voice AI klare Anweisungen geben, z.B. zur Erstellung eines Voice-Over-Skripts für ein LinkedIn Video:

„Du bist eine erfahrene KI im Bereich Voice Content Creation für B2B-Formate auf LinkedIn. Deine Aufgabe ist es, aus einem gelieferten Video und Textbausteinen ein professionelles Voiceover-Skript für LinkedIn-Videos mit einer Länge von 30–90 Sekunden zu erstellen. Du entwickelst dabei den gesprochenen Sprechertext. Die Inhalte richten sich an Fach- und Führungskräfte aus der Büroausstattungsbranche und sollen informativ, klar verständlich und ohne Fremdwörter sein. Dein Stil ist professionell, nahbar und inspirierend. Du formulierst klare und kurze Sätze und achtest auf eine starke Markenwirkung. Optional fügst du am Ende des Sprechertexts einen Call-to-Action hinzu.“

Sobald Sie den generierten Text dann in das entsprechende Voice AI Tool integriert haben, können Sie je nach Software auch weitere Einstellungen über die Funktionen des jeweiligen Tools vornehmen, z.B. zur Aussprache (Schnelligkeit, Betonungen oder Pausen) und Aussehen (Geschlecht, Hautfarbe oder Alter) des Voice Avatars vornehmen. Wie eine solche Umsetzung aussehen kann, zeigt folgendes Beispiel in Murf AI:



Anschließend kann Ihr Voice Avatar in den dafür vorgesehenen Kanälen veröffentlicht werden. So schaffen Sie schnell und einfach ein starkes, konsistentes Audio Branding für Ihre Marke.

In 5 Schritten Ihr Audio Branding gezielt umsetzen

1. Definition Ihrer Markenidentität

Bevor Sie Voice Avatare umsetzen können, ist es essenziell, dass Sie die Kernbestandteile Ihrer Markenidentität definieren:

- **Markenvision und -mission:** Was ist das übergeordnete Ziel und der konkrete Auftrag Ihrer Marke?
- **Markenwerte und -haltung:** Was sind die Grundprinzipien, nach denen Ihre Marke handelt?
- **Markenpersönlichkeit und –kern:** Welche Charakterzüge hat Ihre Marke? Was ist die emotionale Essenz Ihrer Marke?
- **Markenversprechen:** Was ist der konkrete Nutzen, die die Marke Ihren Kunden verspricht?
- **Zielgruppe und Positionierung:** Für wen existiert die Marke? Welche Bedürfnisse adressiert sie? Wie grenzt sie sich vom Wettbewerb ab?

2. Soundstrategie entwickeln

Legen Sie fest, wie Ihre Marke klingen soll. Dazu definieren Sie Ihren individuellen Voice-AI-Styleguide mit den auf Seite 7 erklärten Prompt-Bestandteilen.

3. Voice AI Tools auswählen

Anschließend wählen Sie ein für Ihre Marke passendes Voice AI Tools aus, um damit Ihre Markenstimme KI-generiert zu entwickeln. Um das richtige Tool zu finden, nutzen Sie die Auswahlkriterien von Seite 5.

4. Testen von Prototypen

Mit dem Voice AI Tool generieren Sie anschließend mehrere Varianten Ihrer Markenstimme. Die KI-generierten Voice Avatare testen Sie dann, z.B. indem Sie Kunden in Fokusgruppen zur Wirkung der Markenstimme befragen.

5. Integration in die Touchpoints

Setzen Sie Ihr Audio Branding anschließend in Ihren Kanälen gezielt ein, z.B.:

- **Webseite und App:** z.B. Intro-Sound oder UI-Sounds
- **Social Media:** z.B. Jingles für Reels oder TikToks
- **Telefonie:** z.B. Warteschleifenmusik oder Begrüßungen
- **Events und Messen:** z.B. Soundscapes oder Live-Intros
- **TV- und Vdeo-Werbung:** z.B. Audio-Logos am Ende von Spots

Mit Audio Branding eine starke Markenbindung schaffen

Im digitalen B2B-Marketing eröffnet der Einsatz von Voice AI neue Möglichkeiten für eine klangbasierte Markenführung und macht Audio Branding zu einem zentralen Bestandteil moderner Kommunikationsmaßnahmen. Technologien wie Text-to-Speech und Voice Avatare ermöglichen eine effiziente, konsistente und markengerechte Vertonung von Inhalten.

Ob Social Media Posts, Erklärvideos oder mehrsprachiger Content: Mit Voice AI lassen sich unterschiedlichste Formate schnell und skalierbar vertonen. Besonders wirkungsvoll sind Voice Avatare, die Tonalität, Charakter und Wiedererkennbarkeit einer Marke auditiv transportieren und damit das akustische Pendant zur visuellen Markenidentität bilden. Die Auswahl des passenden Tools ist dabei entscheidend. Während Synthesia insbesondere für videobasierte Markenkommunikation mit KI-Avataren geeignet ist, überzeugt Murf AI durch hochwertige Voiceovers. Lovo AI bietet kreative Möglichkeiten für mehrsprachige Kampagnen mit großer Stimmenvielfalt, während Amazon Polly sich als Lösung für die technische Integration und automatisierte Sprach-ausgabe in Systemlandschaften eignet. Die Wahl sollte sich daher an klar definierten Kriterien wie Stimmqualität, Individualisierung, Nutzerfreundlichkeit und Datenschutz orientieren.

Voice AI ist dabei weit mehr als ein technisches Hilfsmittel, sie ist ein zentrales Instrument für modernes Audio Branding. Unternehmen, die frühzeitig in eine konsistente Audioidentität investieren, stärken ihre Markenbindung, erhöhen ihre Wiedererkennbarkeit und erschließen neue Reichweiten im digitalen Raum.

Julia
Hahn

Prof. Dr.
Markus Gahler

Aivinova GmbH
Wallbergstr. 3
82024 Taufkirchen bei München

Digital Quality GmbH
Gertrudis Str. 4
53909 Zülpich

AIVINOVA

Audio Branding mit KI

AIVINOVA.COM