

KI-Gagenkompass und Vertragsrichtlinien DACH

Herausgegeben vom Verband Deutscher Sprecher:innen e.V.



Dieses Dokument ist aus der Zusammenarbeit der deutschsprachigen Sprecher:innen-Verbände entstanden.

VDS
Deutschland

vds verband
deutscher
sprecher:innen

VOICE
Österreich

VOICE
SPRECHERVERBAND.AT

VSP|ASP
Schweiz

” Vereinigung professioneller Sprecherinnen und Sprecher
Association des speakerines et speakers professionnels
Associazione degli speaker professionisti

*Sprechergagenerhebung & Hinweise im Bereich Synthetische Stimmen
herausgegeben vom Verband Deutscher Sprecher:innen e.V.
auf Basis der globalen Richtlinien*

AI Guidelines for Voice Work

UVA – United Voice Artists

UVA UNITED
VOICE
ARTISTS

Ergänzend zum **VDS-Gagenkompass** bildet der KI-Gagenkompass eine Orientierung für die Berechnung von Sprecher:innengagen in dem neuen Segment der Stimmsynthese (KI-Stimmen).

Inhalt

Präambel

01 Allgemeine Vertragsgrundsätze

02 Stimmsynthetisierung

- 2.1. Neural Learning
- 2.2. Listing Fee
- 2.3. Stimmführung (Speech-to-Speech)

03 Mindestvertonungen live

04 KI-Verwertungsgagen

05 Gebührentransparenz

06 Weitere vertragliche Aspekte

- 6.1. Rechtliche Aspekte
- 6.2. Vertragliche Aspekte

Präambel

"Der Einsatz von KI muss menschliche Entfaltung erweitern und darf sie nicht vermindern. KI darf den Menschen nicht ersetzen."

Alena Buyx, Vorsitzende des Deutschen Ethikrats (2023)

Die internationale Sprechergemeinschaft ist der Überzeugung, dass Werke, die auf menschlichen Emotionen und zwischenmenschlicher Kommunikation beruhen, von echten menschlichen Stimmen vorgetragen werden sollten. Wir sind der Überzeugung, dass nur Menschen in der Lage sind, Gefühle in Worte zu fassen und eine empathische Verbindung mit dem Publikum und den Zuhörer:innen herzustellen. Künstliche Intelligenz (KI) kann die Nuancen, die für eine authentische Ausführung dieser Aufgabe in hoher Qualität notwendig sind, weder meistern noch verstehen.

Unsere Verbände sind sich zugleich darüber im Klaren, dass die Entwicklung synthetischer Stimmen in Kombination mit KI unaufhaltsam voranschreitet und daher eingeordnet und angemessen bewertet werden muss, um allen Marktteilnehmer:innen eine einheitliche Orientierung und eine Berechnungsgrundlage, sowie Schutz vor Missbrauch zu bieten.

Die technischen Möglichkeiten zur Herstellung und Nutzung von KI und synthetischen Stimmen entwickeln sich in immer schnellerem Tempo – zugleich werden die Rechte der Künstler:innen oft übersehen. Vor allem in Europa schützen die DSGVO, Persönlichkeitsrechte, Urheberrechte, Vergütungsansprüche, Haftung usw. die Künstler:innen und ihre Arbeit, während in anderen Ländern der rechtliche Schutz stark variiert.

Die KI-Stimmtechnologie birgt die Gefahr, die wirtschaftliche Grundlage professionell Sprechender im schlimmsten Fall zu vernichten und die Selbstbestimmung über die eigene Stimme zu verlieren. Daher raten die Verbände allen Sprecher:innen weltweit eindringlich, abzuschließende Verträge sorgfältig und gewissenhaft zu prüfen, insbesondere im Hinblick auf den Einsatz und die Anwendung von KI.

Die folgenden Richtlinien und Empfehlungen sollen allen Beteiligten weltweit (z. B. Künstlerinnen, Agenten, Kundinnen, Lizenzgebern, Lizenznehmerinnen, Markenvertretern, Anwältinnen und Gesetzgebern) helfen, vertragliche Vereinbarungen und angemessene Vergütungen für Sprachprojekte mit KI festzulegen. Dieses Dokument dient dabei als Leitfaden für faire Verhandlungen.

01 Allgemeine Vertragsgrundsätze

Um eine faire Zusammenarbeit zu gewährleisten, sollten die folgenden Grundsätze die Basis eines jeden Vertrags bilden.

Generelle Zustimmung

Die Erstellung eines digitalen Abbilds der Stimme einer jeden Person darf nur mit ihrer ausdrücklichen Zustimmung erfolgen. Dabei muss eine spezifische und detaillierte Beschreibung der beabsichtigten Verwendung gegeben werden.

Angepasste Vergütung

Die Vergütung muss den Umfang aller beabsichtigten Nutzungen der digitalen Stimm-Kopie eines Sprechers widerspiegeln und dabei alle Aspekte (quantitativ, zeitlich, räumlich, technisch, medial etc.) berücksichtigen.

Allgemeine Nutzungsbeschränkung

Gänzlich uneingeschränkte Nutzungsrechte zum Einsatz des digitalen Stimm-Duplikats sollten niemals vereinbart werden.

Zeitliche Beschränkung

Zeitlich uneingeschränkte Rechte zum Einsatz des digitalen Stimm-Duplikats sollten niemals vereinbart werden.

Opt-Out-Prinzip

Der Sprecher muss die Möglichkeit haben, die Verwendung seiner Stimme für die KI-Erstellung zum Ende einer jeden Vertragslaufzeit zu beenden. In diesem Fall muss der Lizenznehmer (Auftraggeber) auf Verlangen des Lizenzgebers (Sprecher:in) eine vollständige Löschung der stimmgebenden Daten gewährleisten und darüber hinaus die weitere Herstellung der Stimme oder ihrer Merkmale durch geeignete technische Mittel verhindern.

Stimmtreue (Blended Voices / Morphing)

Im Hinblick auf Rückverfolgbarkeit, Mitbestimmung, Persönlichkeits- und Werbungsrechte, wird das 'Blended Voices / Morphing' von Sprecher:innen entschieden abgelehnt. Wird eine solche Vereinbarung dennoch getroffen, sollte jeder einzelne Sprecher, dessen Stimme einen Teil der gemischten Stimme ausmacht, nicht nur anteilig, sondern vollständig entschädigt werden. Zudem muss er das volle Zustimmungsrecht behalten und sollte unter allen Umständen in der Lage sein, das Repertoire, in dem die gemischte Stimme verwendet wird, zu kontrollieren.

Sprachtreue (Einzigartigkeit der Sprache, Fremdsprachen)

Um internationale Märkte zu schützen und die Kultur der jeweiligen Sprache zu bewahren, sollten KI-generierte Übersetzungen der gesprochenen Sprache nicht vorgenommen werden.

Marktortprinzip

Um Lohndumping durch unterschiedliche Gagenniveaus im internationalen Markt zu vermeiden, sollten sich die Gagen immer mindestens am Standard des Verwertungslandes orientieren.

02 Stimm synthetisierung

Damit KI-Systeme in der Lage sind, künstliche Stimmen zu erzeugen, muss die Originalstimme einer Sprecherin zunächst synthetisiert werden.

Jeder Sprecherin steht es selbstverständlich frei, die eigene Stimme auf eigene Kosten zu synthetisieren und selbstverwaltet anzubieten.

Wird die Stimm synthetisierung jedoch durch einen Kunden / ein Studio / eine Produktion durchgeführt, so muss die Sprecherin dafür vergütet werden.

Zusätzlich ist ein vollumfänglicher Vertrag unter Berücksichtigung sämtlicher Allgemeiner Vertragsgrundsätze (s. Punkt 1) über den Einsatz der Aufnahmen unabdingbar.

Die folgenden Gagen entschädigen nur für den Input der Sprecherin in das KI-System und/oder ihre Studioarbeit, bzw. die Option, die Stimme anzubieten. Sie räumen keinerlei Nutzungsrechte ein. Die Nutzung muss immer zusätzlich vergütet werden.

Genre		Gage			Lizenzen & Anmerkungen
		unterer ø	mittlerer ø	oberer ø	
Stimm synthetisierung	Tagessatz	1.000 €	1.250 €	1.500 €	Lizenzen: Nur interne Nutzung, ausschließlich zum synthetischen Klonen der Stimme und Herstellung einer zu vereinbarenden, marktüblichen Zahl an Demos mit vorab definierten Texten zur öffentlichen Bewerbung der KI-Stimme durch den Vertragspartner, vorausgesetzt, es wurde keine 'Listing Fee' (s. Punkt 2,2) vereinbart. Keine Verwendung zum 'Neural Learning', dem Morphen von Stimmen oder Trainieren von KI-Anwendungen. Ein Aufnahmetag, max. 5 Stunden, inkl. Pausen.
Mindestgage Basis-Synthese		5.000 €	6.250 €	7.500 €	Anwendbar, wenn weniger als 5 Aufnahmetage anfallen. Zum Beispiel wenn bereits Material der Stimme vorliegt.

2.1. Neural Learning

Trainieren von KI-Systemen

Beim 'Neural Learning' wird die Stimmen-Identität einer Sprecherin in das neuronale Netz eines Anbieters eingespeist (z.B. Open AI, Eleven Labs, Microsoft, Google etc.) sodass ihre individuellen Charakterparameter (z.B. Stimmführung, Modulation, Sprachmelodie, Timing, Atmung, Druck, Lebenserfahrung, Persönlichkeit, Charme, Humor uvm.) ein KI-System trainieren und zu einem festen Bestandteil dieses werden. Diese KI kann im Anschluss mit jedem beliebigen Stimmklang, eine Sprachausgabe mit den Charakterparametern dieser Sprecherin generieren oder diese mit anderen Charakterparametern kombinieren.

Das System erfährt mit jeder Einspeisung einen Lernerfolg ('Neural Learning'), **welcher niemals wieder rückgängig gemacht werden kann** - auch wenn die ursprünglichen Trainingsdaten gelöscht werden. Dies

stellt die Basis der Weiterentwicklung der stimmgebenden neuronalen KI-Systeme dar und hat deswegen einen hohen Wert.

Da es bei diesem Lernprozess um mehr als nur das Imitieren des Klangs einer Stimme geht, nämlich um die hochindividuelle und persönliche Stimmidentität, lässt sich eine Abgabe dieser Merkmale mit keiner Vergütung abbilden. Die Sprecherin stellt damit ihre gesamte professionelle Expertise und wirtschaftliche Einzigartigkeit komplett zur Verfügung. Deswegen kann das neuronale Lernen potentiell einen wirtschaftlichen Totalschaden für die Sprecherin bedeuten.

Aus diesem Grunde raten die Verbände allen Sprecher:innen von einer Beteiligung an 'Neural Learning' ab und geben keine Gagen hierfür heraus. Der Wert wäre sonst unermesslich hoch und sollte mindestens im 6-stelligen Bereich liegen.

Wenn dieser Trainingsprozess von Auftraggeberseite nicht eindeutig und nachweisbar ausgeschlossen werden kann, muss die ausdrückliche Zustimmung der Sprecherin zum Neural Learning eingeholt werden und diese dafür durch eine entsprechende 'Neural Learning Fee' vergütet werden, unabhängig davon ob vorab aufgenommenes Material genutzt wird oder gezielt Aufnahmen hergestellt werden. Problematischerweise ist vielen Auftraggebern dabei nicht vollumfänglich bewusst, wie und wo die Trainingsdaten, die sie in ihre Systeme speisen, im Hintergrund weiterverarbeitet werden.

2.2. Listing Fee

Anbieten und Verkaufen von synthetisierten Stimmen

Die Weiterentwicklung der KI-Stimmtechnologie bietet unterschiedlichen Marktteilnehmenden die Möglichkeit, ihre Dienstleistung durch das Feilbieten von synthetisierten Stimmen zu erweitern (z.B. Tonstudios, Sprecherdatenbanken, Agenturen, etc.) Das Angebot synthetisierter Stimmen muss vorab von den ursprünglichen Sprecher:innen autorisiert und sollte, für den Fall, dass das Anbieten und Verkaufen von synthetisierten Stimmen das hauptsächliche oder alleinige Geschäftsmodell darstellt, durch eine 'Listing Fee' vergütet werden, da solche Vermarkter durch die Präsentation von Premium-Stimmen (z.B. bekannte Synchronsprecher:innen, Markenstimmen usw.) ihren Marktwert erheblich steigern können.

Die Listing Fee ist frei zu verhandeln und beginnt, abhängig vom Sprecher, bei symbolischen 150 € pro Jahr und kann bis mehrere Tausend Euro pro Jahr angesetzt werden. Der Betrag ist, nach Vereinbarung, monatlich oder einmal im Jahr im Voraus fällig.

Die Rechte an der digital nachgebildeten Stimme und deren Output verbleiben immer beim Originalsprecher. Die Listing Fee beinhaltet keine Nutzungsrechte. Alle Nutzungsrechte müssen unabhängig von der Listing Fee und der Kompensation der Stimmsynthetisierung verhandelt und jeweils lizenziert werden.

Die Listing Fee sollte zudem immer von einem abzuschließenden Vertrag begleitet werden, der alle Details der Zusammenarbeit, die Vergütung der Nutzungslizenzen und klare Absprachen zur Transparenz, Qualitätskontrolle und Zustimmung rund um den Stimmeneinsatz regelt.

2.3 Stimmführung (Speech-to-Speech)

Führt eine Sprecherin eine oder mehrere fremde/synthetische Stimmen (Stichwort Speech-to-Speech), so muss sie sowohl für den jeweiligen Nutzungsumfang des daraus entstandenen Outputs als auch pro geführte Stimme voll vergütet werden.

Erklärung:

Erst durch eine menschliche, hochwertige Stimmführung wird einer fremden/synthetischen Stimme 'Leben eingehaucht'. Da jede geführte Stimme das Können des Sprechers erneut voll abrufen und die finale Produktion dessen Performance beinhaltet, muss er dafür auch jeweils voll vergütet werden.

03 Mindestvertonungen live

Bei dem Einsatz von KI-Stimmen ist pro Produktion stets eine Basis an originär menschlich eingesprochenem Material herzustellen.

Die deutschsprachigen Verbände sehen die KI-Stimmsynthese nicht als Ersatz der echten Stimme, sondern als Ergänzung dieser. Deshalb, und auch um eine stilistische Vielfaltigkeit in unterschiedlichen Produktionen und für unterschiedliche Marken zu bewahren, sollte bei den meisten Projekten eine Mindestanzahl an Motiven / Sendungen / Kapiteln / Modulen von der echten Stimme vertont und vergütet werden und erst anschließend durch dessen KI-Stimme im gleichen Sprechstil ergänzt werden.

Genre	Mindestvertonungen live	Lizenzen & Anmerkungen
Paid Media	5 - 15 Motive	pro Kampagne / Produkt s. Gagenkompass z.B. TV-Spot, Funk-Spot, Online-Spot usw.
Unpaid Media	1 Motiv / 1 Modul / 2 Min.	pro Kampagne / Produkt s. Gagenkompass z.B. Webvideo, Imagefilm, Präsentation, App usw.
Hörbuch	60 Min. 'Final Audio Hour' (FAH)	pro Hörbuch s. Gagenkompass 1.8 Die Folgestunde gem. Gagenkompass (1.8) wird mit einer Textlänge von 7.200 Wörtern bemessen.
Games	1 Aufnahmestunde	pro Rolle und Titel s. Gagenkompass 1.9 Die Folgestunde gem. Gagenkompass (1.9) wird mit einer Textlänge von 1.200 Wörtern bemessen. Nur anwendbar bei stimmlich beanspruchenden und monotonen Rollen sowie bis zu 10 Pickups. Kein KI-Einsatz bei regulären Rollenaufnahmen.
Synchron	1 Aufnahmestunde	pro Rolle und Titel Nur anwendbar bei bis zu 10 Pickups. Kein KI-Einsatz bei regulären Rollenaufnahmen.
Redaktionelle Inhalte, Dokumentarfilme, TV-Reportagen	15 Min.	pro Film, Filmreihe, Sendung, Titel s. Gagenkompass 1.10

Alle hier nicht aufgeführten Genres benötigen nicht zwangsläufig eine Mindestanzahl an live gesprochenen Anteilen und können, nach Vereinbarung, auch ausschließlich durch die KI-Stimme generiert werden. Die Lizenzen & Anmerkungen zu jedem Genre finden sich im Gagenkompass.

Setzt sich eine Kampagne aus mehreren unterschiedlichen Medien zusammen (z.B. Paid Media & Unpaid Media), so gilt die Mindestanzahl des höherpreisigen Segments (in diesem Beispiel "Paid Media").

04 KI-Verwertungsgagen

In diesem Kapitel geht es um die Nutzung des KI-generierten Outputs. **Grundsätzlich basiert die Vergütung von Sprecher:innen auf dem Erwerb von Nutzungslizenzen ihrer Stimme. Die KI-Technologie ändert an diesem Grundsatz nichts.**

Es gilt:

Nutzung ist Nutzung.

Die technische Anwendung von KI kann Kunden möglicherweise im Produktionsprozess Zeit und Geld einsparen. Die Höhe der Vergütung ist jedoch immer abhängig vom Umfang der Verwertung und der damit erreichten Zielgruppe.

Die Ansprüche aus der Vergabe von Nutzungsrechten für die eigene Stimme bleiben unverändert, unabhängig davon, durch welche Methode diese hergestellt wurden, da der Wert und das Konzept dieser Rechte unverändert bleiben. Diese sind vollumfänglich gem. Gagenkompass zu vergüten.

Mitunter kommt es zu dem Missverständnis, dass nur die hörbare Stimme durch Lizenzen vergütet wird. So ist jedoch sowohl die Stimmfarbe, als auch die Stimmführung Teil der zu vergütenden Leistung. Konkret fällt also bei Nutzung einer Sprachaufnahme als Grundlage für eine andere Stimme (sog. Speech-to-Speech), die volle Gage für die Nutzung beider Stimmen an. Sowohl für die stimmführende Stimme, die evtl. „nur“ die schauspielerische Performance beisteuert, als auch für die am Ende hörbare Stimme, die aus welchen Gründen auch immer darübergerlegt wird.

Auch Paketpreise bei umfangreichen Projekten sind wie bisher individuell zu verhandeln. Hinweise zu den üblichen Standards finden sich ebenfalls im Gagenkompass (Punkt 2).

Die Verwertungsrechte, welche bei KI-Stimmen zur Anwendung kommen, entsprechen den Verwertungsrechten in den jeweiligen Gagenübersichten für das jeweilige Segment.

VDS Deutschland	VDS-Gagenkompass	https://www.sprecherverband.de/service/vds-gagenkompass/
VOICE Österreich	VOICE Honorarsätze	https://sprecherverband.at/de/honorarsaetze
VPS ASP Schweiz	VPS-ASP Tarifliste	https://www.vps-asp.ch/about/tariff

Der Sprecher, der seine Stimme für ein Stimm-Duplikat zur Verfügung stellt, behält die vollständigen Rechte daran und muss entsprechend dem Umfang der Nutzung des digitalen Stimm-Duplikats entschädigt werden.

Angesichts des breiten Spektrums von Anfragen, das von Werbung und System-/Objektstimmen bis hin zu Hörbüchern und darüber hinaus reichen, ist es von entscheidender Bedeutung, die in Punkt 1 dargelegten

allgemeinen Vertragsgrundsätze vollständig zu verstehen und anzuwenden, um eine angemessene und faire Vergütung für die Nutzung zu bestimmen.

Solange die Qualität einer digitalen Replik den Anforderungen eines Kunden an die Darstellung seiner Marke, seines Produkts oder seiner Inhalte mit einer KI-generierten Stimme gerecht wird, gibt es keinen guten Grund, den Lizenzgeber (Sprecher) für diese Verwertung anders zu vergüten als bisher.

05 **Gebührentransparenz**

Vergütungsstrukturen und ihre Bestandteile sind sehr spezifisch und klar in Verträgen zu vereinbaren und zu regeln, um ein Höchstmaß an Transparenz für alle Parteien zu ermöglichen.

Abhängig von der Vereinbarung zwischen dem Sprecher und seiner Kundin und dem Umfang der von dem Sprecher zu erbringenden Leistungen ist die aktuelle Preisliste des Sprechers als fester Bestandteil in den Vertrag aufzunehmen.

Etwaige Mengenrabatte können im Vertrag definiert werden oder projektabhängig im Einzelnen verhandelt werden. Werden Mengenrabatte im Vertrag nicht definiert, so verpflichtet sich der Sprecher, unabhängig vom Umfang des aktuellen Projekts, nicht zur Vergabe von Rabatten und es gilt seine im Vertrag definierte Vergütungsstruktur.

Dem Sprecher sollte das Recht eingeräumt werden, seine vertraglich vereinbarte Vergütungsstruktur in regelmäßigen Abständen (z.B. alle 3 Jahre) an seine aktuelle Gagenliste anzupassen, um etwaige Marktentwicklungen, Inflation, Änderungen des Marktwertes der eigenen Stimme uvm. abbilden zu können. Der Kundin steht es frei, den Vertrag bei einer Preisanpassung seitens des Sprechers zu kündigen / nicht zu verlängern. In diesem Falle ist eine umgehende und vollständige Löschung der stimmgebenden Daten durch die Kundin zu gewährleisten und darüber hinaus die Herstellung der Stimme oder ihrer Merkmale durch geeignete technische Mittel zu verhindern.

06 Weitere vertragliche Aspekte

Bei der Verhandlung eines KI-Vertrags sollten die folgenden Faktoren gründlich bedacht und berücksichtigt werden.

6.1. Rechtliche Aspekte

Persönlichkeits- und Veröffentlichungsrechte - Das Recht an der eigenen Stimme

Die Sprecherin ist durch das Persönlichkeitsrecht, welches das Recht an der eigenen Stimme umfasst, besonders geschützt. Dieses Recht ist nicht übertragbar, auch nicht im Falle einer Stimmvervielfältigung. Die Sprecherin sollte daher die Möglichkeit haben, eine Nutzung der Aufnahmen abzulehnen, welche über den ursprünglich vereinbarten Zweck hinausgeht, auch wenn eine zusätzliche Vergütung angeboten wird.

Urheberrecht und geistiges Eigentum

Es sollte klar kommuniziert und schriftlich festgehalten werden, dass der Sprecher, der die Stimme "leiht", auch die Rechte an der digitalen Kopie besitzt. In der EU sind Sprecher durch das Urheberrecht und geistige Eigentumsrechte geschützt und es können nur Lizenzen für die Nutzung übertragen werden. Alle Beteiligten sollten sich auch darüber im Klaren sein, dass KI-generierte Inhalte nicht automatisch urheberrechtlich geschützt sind.

Datenschutz

Die Lizenznehmerin muss sicherstellen, dass die persönlichen Daten des Sprechers geschützt sind und der Sprecher ordnungsgemäß darüber informiert wird. In der EU ist der Datenschutz durch die Datenschutz-Grundverordnung (DSGVO) streng geregelt. Abgesehen von den eindeutigen Hinweisen wie unter Punkt 1 (Allgemeine Vertragsgrundsätze) dürfen die für die Stimmvervielfältigung erforderlichen Daten und die daraus resultierenden Informationen zu keinem Zeitpunkt außerhalb der EU verarbeitet oder dahin übertragen werden. Das bedeutet, dass sich die Server, auf denen die KI gehostet wird, und die Server, auf denen sich der Algorithmus befindet und das Replikat verarbeitet wird, in der EU (bestenfalls in Deutschland) befinden müssen, um ausreichenden Schutz zu gewährleisten.

Gerichtsstand

Der im Vertrag angegebene Gerichtsstand muss das Land des Wohnsitzes der Sprecherin sein.

Haftungsausschluss

Der Sprecher ist von jeglicher Haftung für die mit seiner digitalen Nachbildung produzierten Inhalte zu befreien. Die gesamte Haftung liegt bei der Auftraggeberin, dem Einsatzbetrieb und dem KI-Anbieter (wie im AI Act definiert). Diese Parteien sollten dafür eine Haftpflichtversicherung abschließen.

6.2. Vertragliche Aspekte

Zustimmung zu neuen Nutzungsformen

Möchte der Kunde das digitale Abbild für ein neues Produkt, Programm oder eine neue Funktion verwenden oder die Nutzung um eine andere, nicht im ursprünglichen Vertrag aufgeführte Weise erweitern, muss er zuvor die erneute und ausdrückliche Zustimmung der Sprecherin in Form einer Vertragsergänzung schriftlich einholen, in der der neue Verwendungszweck schriftlich exakt festgehalten wird.

Exklusivität

Eine etwaige Exklusivität muss genau definiert werden. Jede Form der Exklusivität muss im Honorar berücksichtigt werden. Die entsprechende Vergütung kann durchaus im 6-stelligen Bereich liegen.

Verbot der Namensnennung

In einigen Märkten ist das Recht auf Namensnennung des Sprechers gesetzlich verankert. Wird vertraglich vereinbart, dass der Sprecher nicht genannt wird, muss der Wertverlust des Sprechers gesondert ausgeglichen werden. Dieser kann auf bis zu 100 % des ursprünglichen Honorars geschätzt werden.

Qualität der digitalen Nachbildung

Die Sprecherin hat ein berechtigtes Interesse daran, zu überprüfen, wie gut die digitale Nachbildung ihre natürliche Stimme imitiert und wie gut sie die Sprache umsetzt (z. B. Phonetik, Sprachmelodie, Betonung, Ausdruck). Da eine minderwertige digitale Nachbildung ihrem Ruf schaden könnte, sollte die Sprecherin das Recht haben, das Endprodukt abzunehmen und vor der Veröffentlichung bei Bedarf eine qualitativ hochwertigere oder treffendere Sprachsynthese zu verlangen.

Spezifische Produkte/Programme

Im Vertrag sollte genau festgelegt werden, für welches Produkt, Programm, welche Plattform, welche Rolle, welches Gebiet, welchen Zeitraum und welche Sprache die digitale Nachbildung verwendet werden darf und in welchem Umfang sie den Nutzern zur Verfügung gestellt wird. Eine teilweise unbeschränkte Nutzung kann zwar auch vereinbart werden, allerdings ist eine klare Eingrenzung notwendig, um die Vergütung und die Nutzungsrechte angemessen zu bestimmen.

Ausschluss von bestimmten Inhalten und Missbrauch

Die Sprecherin sollte das Recht haben, bestimmte Arten von Inhalten auszuschließen. Dies kann unter anderem die Verwendung ihrer Stimme für politische, religiöse, erotische/pornografische oder andere kontroverse Inhalte sein. Darüber hinaus sollte jeglicher Missbrauch durch eine der Parteien, die Zugang zu der digitalen Nachbildung haben (Anbieter, Importeur, Vertreiber oder Bereitsteller im Sinne des AI Acts), unterbunden und auf Kosten des Auftraggebers verfolgt werden. Diese Parteien tragen die Sorgfaltspflicht dafür und müssen die Sprecherin unverzüglich informieren, sobald sie von dem Missbrauch oder der Verwendung in ausgeschlossenen Inhalten erfahren.

Klar verständliche Informationen zur Datenverarbeitung

Jede Aufzeichnung, Verwendung oder Speicherung von Sprachdaten ist eine Form der Datenverarbeitung im Sinne der Datenschutz-Grundverordnung (DSGVO), die die menschliche Stimme zu den biometrischen Daten zählt und damit als besonders sensibel einstuft. Daher müssen die Sprecher zum Zeitpunkt der Unterzeichnung der Vereinbarung in einem Datenschutzhinweis ausdrücklich über die betreffende

Datenverarbeitung informiert werden. Außerdem sollte in den meisten Fällen eine ausdrückliche Zustimmung darüber eingeholt werden.

Umgang mit Verschwiegenheitsvereinbarungen (NDAs)

Verschwiegenheitsvereinbarungen sollten immer zeitlich begrenzt sein und niemals dazu verwendet werden, Industriestandards zu untergraben.

Verschwiegenheitsvereinbarungen oder entsprechende Klauseln dürfen niemals leichtfertig unterzeichnet werden. Sie bringen den Sprecher zum Schweigen und schränken seine Möglichkeiten ein, unlauteres Geschäftsgebaren aufzudecken, zu verfolgen und die Umstände der Gemeinschaft, der Öffentlichkeit und der Presse mitzuteilen.

Vertragssprache & geltendes Recht

Der Vertrag sollte in der Muttersprache der Sprecherin verfasst werden, damit diese ihn vollständig verstehen kann - oder zumindest in einer zweisprachigen Version, bei der die Muttersprache der Sprecherin die maßgebliche und rechtlich bindende Sprache ist. Sollte ein Vertrag in einer anderen Sprache (z.B. Englisch) oder ein zweisprachiger Vertrag erforderlich sein, ist es Aufgabe des Auftraggebers / KI-Anbieters, die Kosten für die Übersetzung in die Muttersprache der Sprecherin zu übernehmen. Besondere Aufmerksamkeit gilt es auch dem auf den Vertrag anwendbaren Recht zu widmen, da es für die Auslegung und Durchsetzung der Vertragsbestimmungen entscheidend ist, insbesondere im Hinblick auf Aspekte der Rechteübertragung.

Anwalts- und Gerichtskosten

Dem Sprecher sollte ausreichend Zeit eingeräumt werden, um den Vertrag von einer Anwältin seiner Wahl prüfen zu lassen. Das bedeutet, dass sich die in der Synchron- und Gamingbranche gängige Praxis, nach der Sprecher ihren Vertrag direkt vor oder nach der Aufnahmesession im Studio unterschreiben, im KI-Kontext nicht mehr praktikabel ist. Eine juristische Prüfung und Beratung zu Sprecherverträgen ist heutzutage unerlässlich. Die Kosten hierfür sind vom Auftraggeber zu tragen. In jedem Fall muss jeder Vertrag unter Berücksichtigung der konkreten Umstände der jeweiligen Situation ausgehandelt, angepasst und vervollständigt werden und dann durch "Standard-Klauseln" ergänzt werden.

Haftungsausschluss: Diese Richtlinien und Empfehlungen erheben keinen Anspruch auf Vollständigkeit und können von Zeit zu Zeit geändert werden, da sich die Gesetzgebung in Bezug auf Transparenz von Datenbanken, Kennzeichnung und Nachverfolgbarkeit von KI-Inhalten noch im Entwicklungsstadium befindet und von Land zu Land unterschiedlich sein kann (z. B. der EU-AI Act im Vergleich zu Gesetzen in den USA). Da die Interpretation und Anwendung der Bestimmungen je nach Gerichtsbarkeit und geltendem Recht variieren kann, empfehlen die Verbände, bei Bedarf Rechtsbeistand in Anspruch zu nehmen und die Standards zu überprüfen, die durch die Arbeit nationaler Gewerkschaften und Verbände etabliert wurden.