



DEMO

First chapter only

The Voice Clone Rights Ledger

A Creator's Checklist for AI Voice Contracts, Consent, Usage Limits, and Takedown Evidence

The Voice Clone Rights Ledger

© 2026 Pragma Vision LLC. All rights reserved.

Trademark Notice

Google, Google Pay, Google Cloud, and Android are trademarks of Google LLC. Stripe is a trademark of Stripe, Inc. Cloudflare and Cloudflare Workers are trademarks of Cloudflare, Inc. Supabase is a trademark of Supabase, Inc. OpenAI and ChatGPT are trademarks of OpenAI, Inc. Claude is a trademark of Anthropic, PBC. W3C is a trademark of the World Wide Web Consortium. Visa is a trademark of Visa International Service Association. OWASP is a trademark of the OWASP Foundation. Midjourney is a trademark of Midjourney, Inc. Canva is a trademark of Canva Pty Ltd. Etsy is a trademark of Etsy, Inc. Amazon is a trademark of Amazon.com, Inc. All other trademarks are the property of their respective owners.

No Affiliation

This book is an independent publication. It is not authorized, sponsored, or endorsed by any of the companies or organizations whose products or services are mentioned herein.

No Professional Advice

The information in this book is provided for educational purposes only. It does not constitute legal, financial, investment, tax, or other professional advice. Readers should consult qualified professionals for guidance specific to their situation.

Code Examples

Code examples in this book are provided for illustration only. They may not be suitable for production use without additional validation, error handling, and security review.

Published by Pragma Vision LLC

First edition, 2026.

Contents

1	What Voice Cloning Rights Actually Mean	7
1.1	The Sentence That Used to Be Innocent	8
1.2	The Difference Between “Use My Voice” and “Model My Voice”	9
1.3	What AI Voice Models Actually Do With Your Sample	11
1.4	The Three Use Patterns Buyers Are Actually Building For	12
1.4.1	Pattern A: One-Off Use, No Cloning	12
1.4.2	Pattern B: Voice Cloning for Internal Production Efficiency	12
1.4.3	Pattern C: Full Voice Model Acquisition and Commercial Licensing	13
2	Consent, Usage, Term, Territory, and Sublicensing	15
2.1	The Five Clauses That Decide Everything	16
2.2	Consent: What You Are Actually Saying Yes To	17
2.3	Usage: What Categories of Use Are Permitted	18
2.4	Term: How Long the License Lasts	20
2.5	Territory: Where the Rights Apply	21
2.6	Sublicensing: The Multiplier Clause	21
2.7	The EU AI Act and US Biometric Layer	22
2.7.1	EU AI Act (2024)	22
2.7.2	US Biometric Information Privacy Acts	23
2.8	The Federal Layer: NO FAKES and ELVIS Acts	24
3	Perpetuity and Derivative Voice Red Flags	26
3.1	The Words That Mean More Than They Say	27
3.2	The Perpetuity Family	28

3.3	The Derivative Voice Trap	29
3.4	The Future Technology Clause	30
3.5	The Sneaky Phrases Hidden in Boilerplate	31
3.5.1	The Identity Substitution Pattern	31
3.5.2	The Training Data Pattern	31
3.5.3	The Quality Assurance Pattern	32
3.5.4	The Anonymization Pattern	32
3.6	The Walk-Away Indicators	33
3.7	The Negotiation Frame	34
4	Payment Models—Buyout, Royalty, Approval	36
4.1	The Three Compensation Models	37
4.2	Model 1: Buyout (Lump Sum)	38
4.3	Model 2: Royalty (Per-Use or Revenue Share)	39
4.3.1	Royalty Accounting and Audit Rights	40
4.4	Model 3: Approval Rights	41
4.5	Hybrid Models—The Pattern That Usually Wins	42
4.6	The SAG-AFTRA Reference Rates	42
4.7	What to Avoid: The Below-Scale Bait	44
5	Biometric Data and Model Training Clauses	46
5.1	Why Voice Is Biometric Data	47
5.2	The BIPA Standard	48
5.2.1	What BIPA-Compliant Consent Looks Like	49
5.3	Other US State Biometric Laws	50
5.4	The EU GDPR Article 9 Layer	51
5.5	The Model-Training Consent Clause	52
5.6	The Destruction Attestation	53
6	Takedown Evidence and Monitoring	55
6.1	The Takedown Reality	56
6.2	Tier 1: Audio Watermarking and Provenance Tracking	56

6.2.1	The Provenance Chain	57
6.3	Tier 2: Reverse Audio Search	58
6.3.1	Manual Monitoring Tactics	59
6.4	Tier 3: Platform Takedown Notices	59
6.4.1	DMCA Notice (For Recordings That Reproduce Original Audio)	60
6.4.2	Voice and Likeness Claims (For AI-Generated Audio)	60
6.4.3	Platform-Specific Takedown Procedures	61
6.5	Tier 4: Legal Escalation	62
6.5.1	State Right-of-Publicity Claims	62
6.5.2	Federal Claims	62
6.5.3	Cease and Desist Letters	63
6.6	The Takedown Evidence Pack	63
6.7	The Quarterly Voice-Use Audit	64
7	SAG-AFTRA and Union Considerations	66
7.1	The 2023 Strike and What It Won	67
7.2	The Union Member’s Obligations	68
7.2.1	The Anti-Strikebreaking Rule	68
7.2.2	The Approval Process for Non-Union AI Work	69
7.3	Equity (Theatrical) and Voice Work	70
7.4	NAVA: The Non-Union Reference	70
7.5	The Audiobook-Specific Layer	71
7.5.1	Audible’s ACX AI Position (as of mid-2026)	72
7.6	The Video Game Voice Actor Subspecialty	72
7.7	The Union vs. Non-Union Calculation	73
7.8	The Working Voice Actor’s Decision Framework	74
8	The Voice Rights Ledger	76
8.1	Why a Ledger, Not a File Cabinet	77
8.2	The Eight Core Ledger Columns	77
8.3	Extended Ledger Columns	78

8.4	The Ledger Workflow	79
8.4.1	Daily / Per-Project Entry	79
8.4.2	Monthly Royalty Reconciliation	80
8.4.3	Quarterly Catalog Audit	80
8.4.4	Annual Expiration Review	81
8.5	Tools That Work for the Ledger	81
8.6	Tying the Ledger to Cryptographic Consent	82
8.7	The Long View: A Career-Length System	83
8.8	Starting From Scratch: The 30-Day Setup	84
8.9	The Closing Discipline	85
	What's Next	86
	About Pragma.Vision	88

1

What Voice Cloning Rights Actually Mean

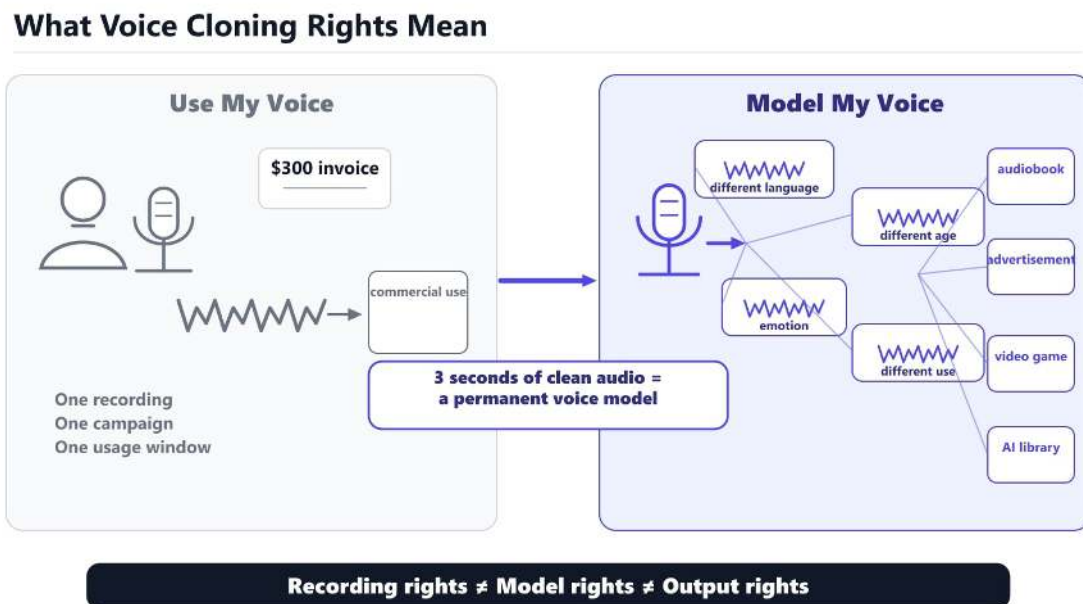


Figure 1. A split-screen comparison separates “Use My Voice” from “Model My Voice”: a \$300 invoice and one commercial-use card sit opposite 3 seconds of clean audio becoming a permanent voice model

1.1 The Sentence That Used to Be Innocent

For most of the history of voice work, the deal looked something like this: a producer hired a voice actor to read a script, paid a session fee, and used the recording in the specific commercial, audiobook, or video for which it was made. If they wanted to use it somewhere else—a sequel, a new market, a different campaign—they came back and paid again. The voice was the actor’s labor for one project, not a permanent rights transfer.

That world ended quietly between 2022 and 2024. The technology to clone a human voice from as little as three seconds of clean audio became commercially available, accurate enough to fool the speaker’s own family, and cheap enough that any production company could afford a license. The economics of voice work flipped: the value of a single recording session was no longer the recording. It was the training data hidden inside the recording. A producer who could extract the voice model from a session could generate unlimited new lines forever, in any language, in any emotional tone, without the actor’s involvement or consent.

3

seconds of clean audio is enough for current-generation AI voice cloning systems to produce convincing synthesis in the speaker’s voice¹

The contracts changed too. Buried in the same boilerplate that used to read “perpetual worldwide license to the recording” is now language that grants the buyer rights to “the voice character,” “derivative voice models,” or “any future technology that may be used to reproduce the speaker’s vocal identity.” Same paragraph structure, completely different meaning. A creator who reads the contract the way they read it in 2018 has already lost.

¹Microsoft Research, “VALL-E: Neural Codec Language Models are Zero-Shot Text to Speech Synthesizers,” 2023; ElevenLabs and Resemble.ai have published similar production benchmarks.

Warning

This book is a checklist for reviewing voice work contracts, capturing consent, and building evidence when your voice is used outside its license. It is not legal advice. Voice rights law is jurisdiction-specific, fast-moving (the EU AI Act came into force in 2024, the NO FAKES Act is pending federal legislation in the US, state biometric laws including Illinois BIPA and Washington's biometric statute differ in scope and penalty structure), and any contract involving meaningful money, exclusivity, or model-training rights deserves review by an entertainment attorney with media-rights experience. The frameworks here generally apply under US law; specific outcomes depend on your state, the buyer's jurisdiction, the technology in question, and the contract language.

1.2 The Difference Between “Use My Voice” and “Model My Voice”

The single most important distinction in any modern voice work contract is between using a recording and modeling the voice. These are not the same right and they should never share the same paragraph in a contract.

Right Granted	What It Permits
Use of the recording	Playing the specific audio file in the specific work it was made for
Use of the recording in derivative works	Editing, syncing, remastering the same audio for related uses
Use of the recording for model training	Feeding the audio into an AI system that produces a voice model
Use of the resulting voice model	Generating new synthetic audio in your voice from text input
Use of the model commercially	Selling, licensing, or sub-licensing synthetic audio of your voice
Sub-licensing the model to third parties	Other companies generating synthetic audio of your voice

A contract that says “buyer receives a perpetual worldwide license to the recording and all derivative works” is not granting model training rights—if the contract is read in the 2018 sense. A contract that says “buyer receives a perpetual worldwide license to the recording, the voice character, and all derivative works in any medium now known or hereafter developed” probably is granting model training rights, depending on how a court interprets “voice character” and “any medium hereafter developed.” The ambiguity is the danger. A producer who later trains an AI model from your recording will argue the contract covered it. You will argue it did not. The contract language at signing time decides who wins.

Key Insight

The cleanest defense against ambiguous voice rights language is to add a specific exclusion clause. A clause that reads “Notwithstanding any other provision of this Agreement, no rights are granted to use the recording, the speaker’s voice, or any data derived from this recording for training, fine-tuning, or development of any artificial intelligence, machine learning, voice synthesis, or voice

cloning system without the speaker’s separate written consent and a separately negotiated fee” closes the ambiguity. Many buyers will agree to this when asked directly. Many will not—which is itself useful information.

1.3 What AI Voice Models Actually Do With Your Sample

Voice cloning systems like ElevenLabs Instant Voice Cloning, Resemble.ai’s Rapid Voice Cloning, Play.ht’s voice cloning, OpenAI’s Voice Engine, Descript Overdub, Speechify’s voice cloning, Murf.ai, and WellSaid Labs all work on roughly the same principle: a deep neural network is trained on a large corpus of human voices to learn how voices generally sound. When you provide a sample of your voice, the system learns the specific characteristics that distinguish your voice from the general population—pitch range, timbre, prosody, accent, breath patterns—and stores them as a parameter set. The sample itself is discarded or retained, depending on the platform. The parameter set is what becomes “your voice” inside the system.

Once the parameter set exists, the system can generate new audio in your voice from arbitrary text. Some systems can also modify the parameter set: making the voice sound younger or older, more excited or more subdued, with different accents, in different languages. The voice model is not just “you reading the same script.” It is a generative engine that can produce anything you might have said but never did.

This matters for contracts because three different things can be sold:

1. The original recording (a fixed audio file)
2. The voice model (the parameter set derived from the recording)
3. The output of the voice model (new audio generated by feeding the model text)

A contract that grants rights to one of these is not granting rights to the others. A contract that is silent on whether the recording can be used to derive a voice model is

ambiguous—and ambiguity in voice rights contracts typically resolves in favor of the larger party with the better lawyers.

1.4 The Three Use Patterns Buyers Are Actually Building For

Producers and studios commissioning voice work in 2026 typically have one of three end uses in mind, even when the contract language obscures it.

1.4.1 Pattern A: One-Off Use, No Cloning

A traditional voice-over job: read a 60-second script for a specific commercial, audio-book chapter, podcast intro, or video narration. The buyer wants the file and standard usage rights. No voice model is built. The session fee is the only payment. This is the world that used to exist and still exists for ethical buyers who do not want the headache of AI rights.

For this pattern, a clean contract with explicit AI exclusion is appropriate. The session fee should match historical rates (\$250–\$500/finished hour for audiobook narration at SAG-AFTRA scale, comparable rates for other use types).

1.4.2 Pattern B: Voice Cloning for Internal Production Efficiency

A studio wants to build a voice model from your sessions so they can generate additional lines without bringing you back. This is increasingly common in animation, video games, and large-scale audiobook production where the producer wants to fix mispronunciations, add new lines, or extend the work without re-booking the original talent.

For this pattern, the appropriate compensation is significantly higher than the session fee—typically 10–15x the base rate for the AI rights, plus a usage-based component

(per-generated-minute fees or per-line fees) and a defined term (3 years, 5 years) after which the rights revert.

1.4.3 Pattern C: Full Voice Model Acquisition and Commercial Licensing

A platform wants to add your voice to their library of available synthetic voices, where other customers can pay to generate audio in your voice for their own projects. This is the “buyout” model promoted by ElevenLabs Voice Library, Murf.ai voice library, and similar marketplaces.

For this pattern, the compensation should reflect the full long-tail value of the voice model. Industry observation suggests \$5,000–\$50,000 for full buyout, depending on the speaker’s commercial profile, with revenue-share alternatives in the 30–50% range of platform-generated revenue from your voice. A flat \$300 “narration fee” that hides a buyout is the trap this book exists to prevent.

Case Study

The \$200 Job That Became a \$50K Model

A working voice actor in Los Angeles accepted what was described as a routine “e-learning narration project”—6 hours of recording, \$200/hour, total \$1,200. The contract referenced “perpetual worldwide license to the recording and derivative works in any medium.” Eight months later, his voice appeared in the buyer’s commercial AI voice library, available to subscribers for \$0.05 per generated minute. Reasonable estimates of the platform’s revenue from his voice over the following year ran into five figures. He had no separate compensation. His attorney’s review of the contract concluded the “derivative works in any medium” language was probably broad enough to cover the model, though litigation could go either way. He settled out of court for a four-figure amount and now adds an explicit AI exclusion clause to every contract he signs. The financial cost of skipping the exclusion clause: roughly an order of magnitude more than the project paid.

Get the complete book — <https://shop.pragma.vision>

DEMO

This is a free preview of the full edition.

Get the complete book at:

<https://shop.pragma.vision>