

**UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK**

PAUL LEHRMAN and LINNEA SAGE, on
behalf of themselves and all others similarly
situated,

Plaintiffs,

v.

LOVO, INC.,

Defendant.

Case No.: 1:24-cv-03770

Hon. J. Paul Oetken

**MEMORANDUM OF LAW IN SUPPORT
OF DEFENDANT'S MOTION TO DISMISS THE SIXTH CAUSE OF ACTION FROM
THE SECOND AMENDED CLASS ACTION COMPLAINT**

TABLE OF CONTENTS

PRELIMINARY STATEMENT 1

STATEMENT OF FACTS 4

ARGUMENT 10

 I. Legal Standard 10

 II. The Training Claim Should Be Dismissed 11

 A. The Training Claim Should Be Dismissed Because It Is Conclusory And
 Speculative 12

 1. Plaintiffs’ “Upon Information and Belief” Pleadings Are Insufficient 12

 2. Plaintiffs Have Not Plausibly Alleged That the Alleged Training
 Infringes Any Sound Recording Copyright in a Non-Conclusory
 Fashion 14

 B. The Training Claim Should Be Dismissed Because Converting a Sound
 Recording to a Written Graph/Chart Is Not Infringement 16

 C. The Training Claim Should Be Dismissed Because the Alleged Copying
 Is Really The Creation of “Output” Which The Court Determined Does
 Not Infringe 17

 D. The Training Claim Should Be Dismissed Because the Alleged Use of
 the Sound Recordings Was Not Substantially Similar and Was
 De Minimis 19

 E. The Training Claim Should Be Dismissed Because Any Alleged Copying
 Is “Fair Use” 22

CONCLUSION 27

TABLE OF AUTHORITIES

Cases	Page(s)
<i>Abdin v. CBS Broad. Inc.</i> , 971 F.3d 57 (2d Cir. 2020)	11
<i>Arista Recs., LLC v. Doe 3</i> , 604 F.3d 110 (2d Cir. 2010)	12
<i>Ashcroft v. Iqbal</i> , 556 U.S. 662 (2009)	10, 13
<i>Attia v. Soc'y of New York Hosp.</i> , 201 F.3d 50 (2d Cir. 1999)	18
<i>Bell Atl. Corp. v. Twombly</i> , 550 U.S. 544 (2007)	9
<i>Bell v. Blaze Magazine</i> , 2001 WL 262718 (S.D.N.Y. Mar. 16, 2001).....	19
<i>Boisson v. Banian, Ltd.</i> , 273 F.3d 262 (2d Cir. 2001)	17
<i>Boyle v. Stephens, Inc.</i> , 1998 WL 80175 (S.D.N.Y. Feb. 25, 1998)	20
<i>Brought to Life Music, Inc. v. MCA Records, Inc.</i> , 2003 WL 296561 (S.D.N.Y. Feb. 11, 2003)	15
<i>Brunswick Recs. Corp. v. Lastrada Ent. Co.</i> , 2023 WL 3010967 (S.D.N.Y. Apr. 20, 2023)	16
<i>Citizens United v. Schneiderman</i> , 882 F.3d 374 (2d Cir. 2018)	12, 13
<i>Cole v. John Wiley & Sons, Inc.</i> , 2012 WL 3133520 (S.D.N.Y. Aug. 1, 2012).....	12, 16
<i>DiTocco v. Riordan</i> , 815 F. Supp. 2d 655 (S.D.N.Y. 2011)	19
<i>Dobson v. NBA Props., Inc.</i> , 1999 WL 97901 (S.D.N.Y. Feb. 24, 1999)	20
<i>Endobotics, LLC v. Fujifilm Healthcare Americas</i> , 2025 WL 1549027 (S.D.N.Y. May 29, 2025)	13
<i>Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., Inc.</i> , 499 U.S. 340 (1991)	11, 19
<i>Gayle v. Allee</i> , 2021 WL 120063 (S.D.N.Y. Jan. 13, 2021)	21
<i>Gottlieb Dev. LLC v. Paramount Pictures Corp.</i> , 590 F. Supp. 2d 625 (S.D.N.Y. 2008)	19, 20, 21, 22
<i>Hamil Am. Inc. v. GFI</i> , 193 F.3d 92 (2d Cir. 1999)	20
<i>In re Jackson</i> , 972 F.3d 25 (2d Cir. 2020)	17
<i>Lehrman v. Lovo, Inc.</i> , 2025 WL 1902547 (S.D.N.Y. July 10, 2025).....	Passim

<i>New York Times Co. v. Microsoft Corp.</i> , 777 F. Supp. 3d 283 (S.D.N.Y. 2025)	20
<i>Ong v. Townsquare Media, Inc.</i> , 2025 WL 2294867 (S.D.N.Y. Aug. 8, 2025).....	21
<i>Peter F. Gaito Architecture, LLC v. Simone Dev. Corp.</i> , 602 F.3d 57 (2d Cir. 2010)	17, 18
<i>Phillips v. Google LLC</i> , 2025 WL 1795894 (S.D.N.Y. June 30, 2025)	15
<i>Ray v. Ray</i> , 799 F. App'x 29 (2d Cir. 2020).....	10, 12
<i>Richardson v. Kharbourch</i> , 2024 WL 50374 (N.D. Ill. Jan. 4, 2024).....	11
<i>Richardson v. Townsquare Media, Inc.</i> , 2025 WL 89191 (S.D.N.Y. Jan. 14, 2025)	22
<i>Rudkowski v. MIC Network, Inc.</i> , 2018 WL 1801307 (S.D.N.Y. Mar. 23, 2018).....	22
<i>Scroggins v. Scroggins</i> , 2017 WL 1047356 (S.D.N.Y. March 16, 2017).....	15
<i>SmartStream Techs., Inc. v. Chambadal</i> , 2018 WL 1870488 (S.D.N.Y. Apr. 16, 2018).....	10
<i>Smith v. Hogan</i> , 794 F.3d 249 (2d Cir. 2015)	10
<i>Staten v. City of New York</i> , 2015 WL 4461688 (S.D.N.Y. July 20, 2015).....	14
<i>Tellabs, Inc. v. Makor Issues & Rights, Ltd.</i> , 551 U.S. 308 (2007)	4
<i>Tongue v. Sanofi</i> , 816 F.3d 199 (2d Cir. 2016)	10
<i>TufAmerica, Inc. v. WB Music Corp.</i> , 67 F.Supp.3d 590 (S.D.N.Y. 2014)	21
<i>Tufenkian Import/Export Ventures, Inc. v. Einstein Moomjy, Inc.</i> , 338 F.3d 127 (2d Cir.2003)	19
<i>Twohig v. Shop-Rite Supermarkets, Inc.</i> , 519 F. Supp. 3d 154 (S.D.N.Y. 2021)	10
<i>VMG Salsoul, LLC v. Ciccone</i> , 824 F.3d 871 (9th Cir. 2016).....	18
<i>Williams v. Calderoni</i> , 2012 WL 691832 (S.D.N.Y. Mar. 1, 2012).....	14
<i>Yamashita v. Scholastic Inc.</i> , 936 F.3d 98 (2d Cir. 2019)	10, 12

Statutes

17 U.S.C. § 106.....	1, 2, 11, 12, 15
17 U.S.C. § 114.....	11, 15

Rules

Federal Rules of Civil Procedure 12(b)(6)..... 1

Secondary Sources

*Scientific Evidence and Forensic Science Since Daubert:
Maine Decides to Sit Out the Dance,*
56 Me. L. Rev. 101 (2004) 16

Defendant LOVO, INC. (“Defendant” or “Lovo”) submits this Memorandum of Law in support of Defendant’s Motion to Dismiss (“Motion”) the copyright infringement training claim (the Sixth Cause of Action) from the Second Amended Class Action Complaint (“Cplt.” or “Complaint”) pursuant to Federal Rules of Civil Procedure 12(b)(6).

PRELIMINARY STATEMENT

In its July 10 order, this Court dismissed most of Plaintiffs’ claims with prejudice including all trademark claims and most copyright claims. The Court permitted Plaintiffs to try to replead their “training-based infringement claim” using information and belief pleading. Plaintiffs have tried, but their attempt fails. Their Training Claim (defined below) remains conclusory, speculative, and factually unmoored from generally available information or otherwise. While providing some detail about a few models for artificial intelligence generally neither those models nor the blog posts upon which Plaintiffs claim their allegations are based describe any copying or duplicating of actual sounds from original sound recordings or any indication Lovo used these models. Instead, the key allegations of “copying” for training do nothing more than allege conclusory words from Section 106 that, even as used, do not articulate any violation. The basic point seems to be that Plaintiffs think Lovo converted their sound recordings into spectrograms, which are visual representations (charts or graphs) showing attributes of sound data, and *this* constitutes copying or duplicating actual sounds. This is wrong and has no legal basis. The training claim should be dismissed, this time with prejudice, for the following reasons:

First, Plaintiffs Training Claim remains conclusory and speculative. The various “upon information and belief” allegations are conclusory, lack sufficient factual bases, and are facially speculative. Even allegations upon information and belief must be more than unsubstantiated suspicions and baseless assumptions. The two blog posts Plaintiffs cite do not support the use of

sound recordings Plaintiffs describe, or even state or imply that Lovo uses those processes. The conclusory mere invocation of words from 17 U.S.C. § 106(1) such as “duplicate,” “rearrange,” “remix,” or “fixed” is, at best, a threadbare recitation of legal conclusions with no support. In fact, Plaintiffs do not even make those conclusory assertions in a manner that in fact describes behavior that *would* infringe a sound recording copyright even in a conclusory fashion. Plaintiffs do not even identify the type of copyright at issue (although it is established that it is a copyright for a sound recording), which Section 106 right Lovo allegedly infringed, or what aspect of the sound recordings is allegedly protectible (and the Court has already determined that “voices *qua* voices” are not protected) and what protected part was allegedly copied.

Second, Plaintiffs Training Claim is premised on the conversion of all the sound recordings into what Plaintiffs call “spectrogram representations” which are visual graph or chart like representations of sound data from the sound recordings. It is clear that these spectrograms are not “actual sounds” which are “fixed in the sound recordings”, and, therefore, creation of these visual representations cannot infringe a sound recording copyright. Thus, all the alleged subsequent copying, remixing, rearranging, or creation of a sound from these spectrograms cannot be activity which infringes a sound recording copyright. There is no other alleged copying.

Third, the Training Claim is based upon creation of the spectrogram and the subsequent creation of phonemes, nascent clones, and other recordings from those spectrograms. But, this is precisely the type of claim this Court has already dismissed (the “output” claim); there is no copyright protection for “voices *qua* voices” and “imperfect mimicry”. Thus, the Training Claim is just a repackaged version of the output claim and should be dismissed for the same reason.

Fourth, the Training Claim does not plausibly allege “substantial similarity” which requires an allegation both that “protected expression” was copied from the sound recordings and

that the amount of protected expression copied was “more than *de minimis*”. Plaintiffs have not plausibly alleged either prong, or that conversion to a spectrogram included any protected attributes. In fact, the only attributes clearly captured in a spectrogram are those that are included in what the Court has already found not protected (*i.e.*, “pitch, loudness, tone, timbre, cadence, inflection, breathiness, roughness, strain, jitter, (variation in pitch), shimmer (variation in amplitude), spectral tilt and overall intelligibility”). Any nascent clones or phonemes allegedly created from the spectrograms would have the same limitation. There is no alleged protected expression from the sound recordings. Even if there were some alleged protected material, there is no allegation that any such material was copied or that there is more than a *de minimis* amount of such putative protected material being copied in any individual copy. The small units of sound processed clearly are not more than *de minimis*.

Finally, the Training Claim pleads a clear case of “fair use” and all four factors favor that result. First, the purpose and character of the use is clearly transformative as (a) Lehrman and Sage allege that they did not intend these recordings for training artificial intelligence or for any commercial use; (b) training artificial intelligence is quintessentially transformative as the use does not try to replace or supplant the original sound recordings; (c) there cannot be any infringing output to the public, as the Court has already dismissed the infringement output claim; and (d) the alleged copying literally transformed sounds into visual representations of attributes. Second, the “nature of the copyrighted work” slightly favors fair use for, as alleged, there is little if any expressive or creative content in these sound recordings created for Lovo with Lovo’s script solely for non-commercial academic or internal use, even if allegedly unpublished. Third, the amount and substantiality of the portion copied made accessible to the public strongly favors fair use because here, as alleged, no alleged infringing copy will be made accessible to the public as the

Court dismissed the output claim. Fourth, as alleged, there will be no effect of the allegedly infringing use for training upon the potential market for *these sound recordings* because, as alleged, there was no market for these sound recordings; they were allegedly made for Lovo, with Lovo's scripts, for Lovo's internal or academic use.

For each or all of these reasons, the Court should dismiss the Sixth Cause of Action (the Training Claim) with prejudice.

STATEMENT OF FACTS

Solely for this motion, Lovo accepts as true the Complaint's allegations as supplemented by documents incorporated or referenced therein. *Tellabs, Inc. v. Makor Issues & Rights, Ltd.*, 551 U.S. 308, 322 (2007) (courts "accept all factual allegations in the complaint as true" and may consider "documents incorporated into the complaint by reference, and matters of which a court may take judicial notice.").

Lovo is a technology company with proprietary software driven by artificial intelligence ("AI") that allows clients to create and edit voice-overs (Cplt. ¶5). Lovo sells a text-to-speech service that allows its clients to generate voice-over narrations (*id.* ¶20). It uses "its AI-driven software known as 'Generator' or 'Genny,' and generate a professional-quality voice-over based on certain criteria ... customers can choose between ... male or female voices, regional accents, and older or younger-sounding voices" (*id.*). "Genny is LOVO's artificial intelligence voice cloning product that creates a digital copy of a human voice using AI and machine learning." (*id.* ¶119). Lovo's technology "turns written text into speech, providing you with a wealth of human-like voices to bring your content to life" (*id.* ¶139 (screen shot)). "Genny by LOVO AI also includes voice editing functions, allowing you to alter the emphasis, pronunciation, cadence, and speed of a voice." (*id.*). Genny "was trained using data from thousands of voices and thousands of hours

of recordings” (*id.* ¶171). Thus, the voices Genny creates are “AI-generated” and not recordings of anything actually said by anyone (*e.g., id.* ¶¶162, 219). These AI-generated voices are also called “clones” (*id.* ¶¶22, 37, 129)

Plaintiffs Mr. Lehrman (“Lehrman”) and Ms. Sage (“Sage”) allege that, in May 2020, Lehrman was “contacted via the Fiverr website by someone identified as User25199087” to sell voice recordings (*id.* ¶44). Fiverr.com is a website that “connects freelancers with clients looking for digital and creative services” (*id.* ¶43). Plaintiffs claim Lehrman communicated with “User25199087” via Fiverr without knowing his identity (*id.* ¶44). After Lehrman was already “working through” – i.e., recording – the “scripts” (*id.* ¶49), “User25199087” told Lehrman informally via Fiverr (a) “Your voiceover will be used for academic research purposes only (*id.* ¶48); (b) “The scripts will not be used for anything else” (*id.* ¶50); and (c) “The script and your finished file will be used for research purposes only.” (*id.* ¶52). Lehrman delivered the recordings via Fiverr and was paid \$1,200 (*id.* ¶¶53-54). Plaintiffs allege Lehrman made 104 recordings of scripts provided by User25199087 (*id.* ¶53) and that he registered certain audio files with the copyright office in August of 2024 (*id.* ¶55).

Plaintiffs allege Lehrman’s voice was being used by an “AI bot” on a podcast (Cplt. ¶¶63-64). But, Plaintiffs acknowledge Lehrman never recorded this; it was either AI-generated or from an AI bot (*id.*). Plaintiffs also claim (a) Lehrman’s “voice” was used to promote Lovo’s “Genny Service” (*id.* ¶105); (b) his “voice” was used “as part of the subscription service under the stage name ‘Kyle Snow’” (*id.* ¶¶ 107); and (c) Lovo had an article on its website that “featured” and “promoted Mr. Lehrman’s voice” as Kyle Snow (*id.* ¶ 139). But Plaintiffs do not allege any of the above were recordings Lehrman sold to Lovo; rather, what they call “his voice” is an AI generated voice that sounds like his voice (*id.* ¶140).

Plaintiffs also allege on October 29, 2019, Sage “received a message on the Fiverr platform from ‘tomlsg’ offering her a contract to produce and record test scripts for radio ads” (*id.* ¶75). Sage also received a message “from ‘tomlsg’ at [the] anonymous company offering the job” telling her these “test scripts” would not be “disclosed externally” and “will only be consumed internally, so will not require rights of any sort” (*id.* ¶77). Plaintiffs claim Sage later discovered “that LOVO had been using, manipulating, and editing her voice in multiple promotional and commercial ways, for many years, without her permission” (*id.* ¶84). Plaintiffs allege (a) “in 2020, Lovo used both a recording of her actual voice – ... and a cloned version of her voice with the identical content ... at the Berkeley SkyDeck Demo Day Spring 2020 event”; and (b) Lovo “posted the presentation publicly to YouTube” (*id.* ¶¶85-86); (c) “Charlie Choi also played Ms. Sage’s recordings at the 2020 CES Tech Conference” (*id.* ¶91); and (d) “Lovo made a cloned version of Ms. Sage’s voice available to Lovo subscribers under the name ‘Sally Coleman’” (*id.* ¶92). Thus, as alleged, the Sally Coleman voice is AI generated and not Sage’s actual recorded voice. Plaintiffs allege Sage registered the “audio file” she sold to Lovo with the Copyright Office and received a registration number for “10 audio tracks” with a June 18, 2024 registration date (*id.* ¶82). The Copyright Office shows these tracks and those registered by Lehrman were registered as sound recordings (Berg. Ex. 1-3).¹ As alleged, the AI generated voices called Kyle Snow and Sally Coleman are both generated by Genny (*e.g.*, Cplt. ¶¶20, 37, 141, 233).

Plaintiffs allege “upon information and belief” that (a) Genny was “pretrained on sound recordings including the Copyrighted Works to recognize certain patterns and characteristics of the human voice.” (Cplt. ¶120); (b) “LOVO processed this dataset of sound recordings into spectrogram representations ... **which are visual representations of audio that show how the**

¹ Declaration of William Bergesch (“Berg.”) with, *inter alia*, exhibits with screenshots of the registration information.

energy of different frequencies changes over time” (*id.* ¶121) (emphasis added), which included “breaking down the data ‘into individual soundwaves so the AI can understand it,’ then ‘labels these sound waves and their corresponding phoneme, the smallest unit of a sound in language’” (*id.*); (c) “LOVO trained the Kyle Snow and Sally Coleman voices **by converting the Copyrighted Works into spectrogram representations that duplicated every aspect of them.**” (*id.* ¶122) (emphasis added); (d) once the recordings were converted to spectrogram representations, LOVO “used one of two processes to train Genny, Kyle Snow, and Sally Coleman to sound like human voices ... It either used a generative adversarial network (GAN) or a process known as likelihood training.”; and (e) either way, “LOVO’s training involved the repeated duplication, rearranging, recapturing, and remixing of the copyrighted sound recordings.” (*id.* ¶123).

Plaintiffs allege that (a) “[a] GAN is a machine learning process comprising two models: one model attempts to generate audio that sounds human and the other guesses which piece of audio is human and which is computer-generated.” (Cplt. ¶124); (b) “Likelihood training operates similarly but does not learn by computer generated example. Instead, the model is trained by reference to the voice it is attempting to mimic in spectrogram form” (*id.* ¶125); and (c) both models “involve constant engagement with the sound recordings of the voice that the model is learning to emulate” (*id.* ¶126).

Plaintiffs then allege upon information and belief that (a) “LOVO trained Genny, Kyle Snow, and Sally Coleman either by comparing the Copyrighted Works to the nascent clones of Mr. Lehrman’s and Ms. Sage’s voices using a GAN, or learning to predict patterns by reference to the voice it is attempting to mimic in spectrogram form using the likelihood method.” (*id.* ¶127); (b) “the Kyle Snow and Sally Coleman voices fuse the knowledge that the Genny model learned

in both the pretraining and training phases with the particular aspects of Mr. Lehrman and Ms. Sage’s voices exhibited in the Copyrighted Works”; (c) “the training process necessarily involved the duplication, rearranging, recapturing, and remixing of the copyrighted sound recordings since the only way to program a computerized system to replicate a human voice is by giving that system examples of the human voice to be emulated.” (*id.* ¶133); and (d) this “training” infringed their copyrights by “processing” which purportedly “involved the repeated duplication of the actual sounds fixed in the Copyrighted Works” (*id.* ¶254) and creating a purported “neural network” for each plaintiff which included “rearranged, remixed, and repurposed versions of the sounds fixed to the Copyrighted Works” (*id.* ¶¶255-256).

Plaintiffs cite only two purported sources for this information: (a) Carol Moh, AI Voice Cloning: What It Is and How It Works, <https://perma.cc/R7JS-9BW2> (Cplt. ¶121); and (b) Carol Moh, Generative AI: Everything You Need to Know, <https://perma.cc/5WPV-Q9GL>, (*id.* ¶124) (printed copies are attached at Berg., Ex. 4-5). Yet neither source discusses use of original recordings as described in the Complaint or the method used by LOVO.

The Original Complaint was filed on May 16, 2024 (ECF 1). Defendant filed a first motion to dismiss on July 29 (ECF 16-17). In response, on August 7, Plaintiffs requested to file an amended complaint (ECF 18). The Court granted that request (ECF 20). Plaintiffs filed the First Amended Complaint on September 25 (ECF 22). Defendant filed a motion to dismiss on November 25, 2024 (ECF 27-29), which was opposed on January 10, 2025 (ECF 33), with a reply on January 31, 2025 (ECF 39). On July 10, 2025, the Court granted the motion in part, dismissing most of the claims with prejudice (ECF 45, and *Lehrman v. Lovo, Inc.*, 2025 WL 1902547 (S.D.N.Y. July 10, 2025)). Specifically, the Court held that “for the most part, Plaintiffs have not stated cognizable claims under federal trademark and copyright law”. *Id.* at *1. The Court also

dismissed the fraud, conversion, unjust enrichment, unfair competition, and all claims for “John Doe” and the “consumer class”. *Id.* at *27-28, n. 30. In dismissing the “output” copyright infringement claim, the Court explained that since Genny creates “new recordings that mimic attributes of their voices”, there can be no copyright infringement on this output because “[c]opyright protection does not extend to this kind of imperfect mimicry, even when accomplished using advanced technology”. *Id.* at *20. The Court denied the motion to dismiss the breach of contract claim and claims pursuant to Sections 50/51 of the New York Civil Rights Law and Sections 349/350 of the GBL, and the infringement claim based on use of the actual recordings (“direct”). *Id.* at *24-27.

The court permitted repleading only with regard to Plaintiffs’ training copyright infringement claim where the Court dismissed the claim, noting that Plaintiffs did not explain “what training is or how it works, even at a very high level of generality”. *Lovo*, 2025 WL 1902547, at *18. The Court suggested that Plaintiffs could try to replead using “information and belief” allegations based upon “generally available information”. *Id.* at *19. The Court also directed that “[i]f Plaintiffs do elect to amend, Defendants shall file any motion to dismiss the amended claims within 21 days of Plaintiffs’ amended complaint. Defendants’ answer to the other remaining claims shall not be due until 14 days after the Court resolves any such motion to dismiss.” *Id.* at *29.

Plaintiffs amended their complaint on July 31, 2025 (ECF 47). The Second Amended Complaint, *inter alia*, included an attempt to replead an infringement training claim (*e.g.*, Cplt. ¶¶119-134; 249-262) (the “Training Claim”). The Court granted a requested schedule for the motion to dismiss of September 15 for the motion, October 6 for the opposition and October 31 for the reply (ECF 48-49). The Court also later granted a joint motion to extend the end of fact discovery to January 20, 2026. (ECF 50-51). In accordance with these orders, Lovo is filing this

motion and will answer any remaining claims 14 days after the Court resolves this motion.

ARGUMENT

I. Legal Standard

A complaint must “state a claim to relief that is plausible on its face.” *Bell Atl. Corp. v. Twombly*, 550 U.S. 544, 570 (2007). “Plausible” allegations allow the court to reasonably infer “that the defendant is liable for the misconduct alleged.” *Ashcroft v. Iqbal*, 556 U.S. 662, 678 (2009) (citation omitted). Mere “labels and conclusions, and a formulaic recitation of the elements of a cause of action” fail. *Twohig v. Shop-Rite Supermarkets, Inc.*, 519 F. Supp. 3d 154, 159-60 (S.D.N.Y. 2021) (quotations omitted). “[A] court may consider the full text of documents that are quoted in or attached to the [complaint], or that the [plaintiff] either possessed or knew about ... in bringing the claims.” *SmartStream Techs., Inc. v. Chambadal*, 2018 WL 1870488, *3 (S.D.N.Y. Apr. 16, 2018). A document referenced in a complaint “control[s] and this Court need not accept as true the allegations” in the complaint. *Tongue v. Sanofi*, 816 F.3d 199, 206 n.6 (2d Cir. 2016) (citation omitted). But, affidavits/declarations supplementing pleadings cannot be considered at all as they are not “legal instruments.” *See, e.g., Smith v. Hogan*, 794 F.3d 249, 254 (2d Cir. 2015). Thus, the Court cannot consider Plaintiffs’ unsworn declarations (ECF 47-1; Cplt. ¶73). Also, even where the facts are “peculiarly within [Defendant’s] control”, Plaintiff, even if pleading upon information and belief, must “marshal more than unsubstantiated suspicions”. *Yamashita v. Scholastic Inc.*, 936 F.3d 98, 107 (2d Cir. 2019); *see also Ray v. Ray*, 799 F. App’x 29, 31 n. 2 (2d Cir. 2020) (no “carte blanche to make baseless assumptions about otherwise permissible conduct”).

II. The Training Claim Should Be Dismissed

Plaintiffs have attempted to replead the Training Claim based upon “information and belief” but their attempt still fails to state a claim. As the Court noted, the relevant copyright at issue here is for a sound recording. *Lovo*, 2025 WL 1902547, at *17; Berg. Ex. 1-3. Statutorily, the exclusive rights for a sound recording are limited to “the right to duplicate the sound recording in the form of phonorecords or copies that directly or indirectly recapture the actual sounds fixed in the recording,” “the right to prepare a derivative work in which the actual sounds fixed in the sound recording are rearranged, remixed, or otherwise altered in sequence or quality,” the right “to distribute copies or phonorecords of the copyrighted work,” and the right “to perform the copyrighted work publicly by means of a digital audio transmission.” 17 U.S.C. § 114 (a), (b); 17 U.S.C. § 106(1), (2), (3), (6). “The rights of duplication and derivation, with respect to sound recordings, ‘do not extend to the making or duplication of another sound recording that consists entirely of an independent fixation of other sounds, even though such sounds imitate or simulate those in the copyrighted sound recording.’” *Lovo*, 2025 WL 1902547, at *17 (quoting 17 U.S.C. § 114(b)). Thus, “[t]he right in sound recordings are limited to direct duplication of the *actual sounds affixed in the recording*, in other words, copying of the actual recording itself”. *Richardson v. Kharbouch*, 2024 WL 50374, at *8 (N.D. Ill. Jan. 4, 2024) (emphasis in original).

The elements of a direct infringement claim are: “(1) ownership of a valid copyright, and (2) copying of constituent elements of the work that are original.” *Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., Inc.*, 499 U.S. 340, 361 (1991). “To satisfy the second element, a plaintiff must demonstrate that: (1) the defendant has actually copied the plaintiff's work; and (2) the copying is illegal because a substantial similarity exists between the defendant's work and the protectible elements of plaintiff's work.” *Abdin v. CBS Broad. Inc.*, 971 F.3d 57, 66 (2d Cir. 2020) (citation omitted). “The word ‘copying’ is shorthand for the infringing of any of the copyright owner's ...

exclusive rights ... described in § 106.” *Arista Recs., LLC v. Doe 3*, 604 F.3d 110, 117 (2d Cir. 2010). A plaintiff must allege “(1) which specific original works are the subject of the copyright claim, (2) that plaintiff owns the copyrights in those works, (3) that the copyrights have been registered in accordance with the statute, and (4) by what acts during what time the defendant infringed the copyright”. *Cole v. John Wiley & Sons, Inc.*, 2012 WL 3133520, at *11 (S.D.N.Y. Aug. 1, 2012) (quotation omitted) (claim dismissed).

A. The Training Claim Should Be Dismissed Because It Is Conclusory And Speculative

1. Plaintiffs’ “Upon Information and Belief” Pleadings Are Insufficient

The allegations of the Training Claim remain general, conclusory, and speculative, and are not saved by the invocation of “upon information and belief.” The Second Circuit has explained that even where the facts are “peculiarly within [Defendant’s] control,” Plaintiff, even if pleading upon information and belief, “must marshal more than unsubstantiated suspicions”. *Yamashita*, 936 F.3d at 107; *see also Ray*, 799 F. App’x at 31 n. 2 (even where “necessary facts ‘are peculiarly within the possession and control of the defendant,’” Plaintiff does not have “carte blanche to make baseless assumptions about otherwise permissible conduct.”); *Citizens United v. Schneiderman*, 882 F.3d 374, 384 (2d Cir. 2018) (affirming dismissal under 12(b)(6); “upon information and belief” are not “magic words,” and “[a] litigant cannot merely plop ‘upon information and belief’ in front of a conclusory allegation and thereby render it non-conclusory.”).

Here, Plaintiffs have added to the Complaint formulaic recitations of the statutory language lacking factual basis or enhancement. They have talismanically alleged rearranging, recapturing, remixing (i.e., some statutory words) of recordings and the phrase sounds fixed in the sound recordings without any factual basis or enhancement or detail (Cplt.¶¶ 123, 126, 133, 254, 255). This amounts to nothing more than a threadbare recitation of the statutory elements. “A pleading

that offers ‘labels and conclusions’ or ‘a formulaic recitation of the elements of a cause of action will not do.’ Nor does a complaint suffice if it tenders ‘naked assertion[s]’ devoid of ‘further factual enhancement.’ *Iqbal*, 556 U.S. at 678; *see also Endobotics, LLC v. Fujifilm Healthcare Americas*, 2025 WL 1549027, at *2 (S.D.N.Y. May 29, 2025) (dismissing claim noting that “a plaintiff ‘cannot merely plop ‘upon information and belief’ in front of a conclusory allegation and thereby render it nonconclusory”) (quoting *Citizens United*, 882 F.3d at 385).

Plaintiffs try to bolster their pleading by citing two blog posts as if information about using sound recordings in training AI are found therein, but neither of those posts actually discusses copying, rearranging, recapturing, or remixing parts of any original sound recordings. Berg. Exs. 4-5. One of the blog posts does explain that when cloning a particular voice from the perspective of the AI, the original sound recordings are “broken down into individual soundwaves so the AI can understand it. The AI then labels these sound waves with their corresponding phoneme, the smallest unit of a sound in language.” Berg. Ex. 4, at 8-9. But, this does not involve any remixing, copying, rearranging, or duplicating of original actual sounds. Moreover, there is no alleged basis to think that Lovo used the process described in the Training Claim for Genny or otherwise. Nor is there any factual bases presented that any available named Lovo AI generated voice (i.e., by Genny) were specifically created using only one individual’s recordings. As Plaintiffs allege, Genny “was trained using data from thousands of voices and thousands of hours of recordings” (Cplt. ¶171).

Further, Plaintiffs plead alternative and inconsistent “facts” about what method Lovo purportedly used to “train Genny, Kyle Snow, and Sally Coleman”. In both paragraph 123 and 129, Plaintiffs allege that Lovo used one of two methods, the generative adversarial network (GAN) method or what they call “likelihood training”. These two methods, as described, are

inconsistent and the use of an either/or alternative pleading highlights its speculative nature. Plaintiffs have no alleged basis to claim that Lovo used either method and, as noted, there is no factual predicate in either blog post for any claims of infringing behavior including improper copying, rearranging, recapturing, or remixing here. This is insufficient. “Information and belief” pleadings must be “accompanied by a statement of the facts upon which the belief is founded”, and remain insufficient where not accompanied by pleading of known information sufficient to render them “anything more than speculative claims or conclusory assertions.” *Staten v. City of New York*, 2015 WL 4461688, at *5 (S.D.N.Y. July 20, 2015), *aff’d*, 653 F. App’x 78 (2d Cir. 2016); *see also Williams v. Calderoni*, 2012 WL 691832, at *7 (S.D.N.Y. Mar. 1, 2012) (upon information and belief pleadings insufficient where plaintiff pointed to no information that would render his statements anything more than speculative claims or conclusory assertions). The same is true here, and Plaintiffs’ allegations confirm they have no factual bases for the speculative allegations of wrongdoing by Lovo in the Training Claim. As such, the Training Claim should be dismissed.

2. Plaintiffs Have Not Plausibly Alleged That the Alleged Training Infringes Any Sound Recording Copyright in a Non-Conclusory Fashion

The Training Claim as a whole is also general and conclusory even if crediting the upon information and belief allegations. Plaintiffs do not explicitly allege what type of copyrights were obtained (the Copyright Office documents show they are “sound recordings” (Berg. Ex. 1-3) as previously recognized by the Court). Nor do they plead what aspect of the recordings are “protectible” (which cannot include “attributes of their voices like ‘pitch, loudness, tone, timbre, cadence, inflection, breathiness, roughness, strain, jitter, (variation in pitch), shimmer (variation in amplitude), spectral tilt, and overall intelligibility, and other qualities” or their “voices *qua* voices” *Lovo*, 2025 WL 1902547, at *17, 20).

They also do not identify what purported “sound recording” copyright rights are being infringed. They do not allege which sections 17 U.S.C. § 106(1), (2), (3), or (6) are the basis for the Training Claim. While creating a word salad of conclusory terms found in some of these provisions (i.e., “fixed”, “sound recording”, “rearranged”, “remixed”, “derivative”, and “duplicated”), Plaintiffs neither utilize them to allege specific infringing behavior, nor put them together to allege that Lovo “duplicate[d] the sound recording in the form of phonorecords or copies that directly or indirectly recapture[d] the actual sounds fixed in the recording,” or that Lovo prepared an original “derivative work in which the actual sounds fixed in the sound recording” are arranged, which meets the definition of a derivative work. 17 U.S.C. § 114 and § 106(1), (2), (3), (6). Plaintiffs do not and cannot allege there was any distribution or performance of the sound recordings as part of the training. Nor do (or could) they claim that any artificially generated snippet of/from the sound recordings infringes because the infringement “output” claims were dismissed. *Lovo*, 2025 WL 1902546, at *20. This failure is fatal to Plaintiffs claim as they simply do not plausibly assert that the alleged training violated any cognizable sound recording rights that Plaintiffs possess. *See, e.g., Brought to Life Music, Inc. v. MCA Records, Inc.*, 2003 WL 296561, at *1 (S.D.N.Y. Feb. 11, 2003) (no allegation defendant “(i) reproduced [her] work; (ii) prepared derivative works based on [her] work; (iii) distributed copies of [her] work to the public for sale; or (iv) performed [her] work publicly”); *Scroggins v. Scroggins*, 2017 WL 1047356, at *9 (S.D.N.Y. March 16, 2017) (“a plaintiff must allege, *inter alia*, “particular infringing acts” by that defendant, and the specific rights that such acts violated”); *Phillips v. Google LLC*, 2025 WL 1795894, at *5 (S.D.N.Y. June 30, 2025) (“Plaintiff must allege specific unauthorized acts by Defendants, such as reproduction, distribution, or public performance of his musical works.”); *Brunswick Recs. Corp. v. Lastrada Ent. Co.*, 2023 WL 3010967, at *4 (S.D.N.Y. Apr. 20, 2023)

(“a copyright complaint must describe activity that could constitute infringement”); *Cole*, 2012 WL 3133520, at *12-13 (“vague and expansive” allegations fail where lacking detail as to “by what acts” infringement occurred).

B. The Training Claim Should Be Dismissed Because Converting a Sound Recording to a Written Graph/Chart Is Not Infringement

Even if the Court determines the Training Claim is not speculative and conclusory, the claim still fails because Plaintiffs allege that the critical step for Lovo to “train” anything and **the basis for all subsequent steps** is the purported conversion of all sound recordings into what Plaintiffs call “spectrogram representations” or “spectrograms” (Cplt. ¶¶121, 122, 125, 129) which they allege are a “visual representations of audio that show how the energy of different frequencies changes over time mapped to a particular scale designed to match how humans perceive sound” (*id.* ¶121). They further allege that this process “break[s] down the data ‘into individual soundwaves so the AI can understand it,’ then ‘labels these sound waves and their corresponding phoneme, the smallest unit of a sound in language’” (*id.*). In other words, the sound recording is converted into a chart or graph showing various sound data.² All subsequent alleged training and generation of clones or sounds, including any alleged duplication, remixing, rearranging, recapturing, and reprocessing, are created using these spectrograms and not the “actual sounds fixed in the sound recordings.”

Conversion of a sound recording to a graphic or visual form does not infringe on any protected rights here because visual or graphic representations are not “actual sounds” nor “fixation” of sounds themselves, with the meaning of the Copyright Act. Also, a subsequent,

² A “voice spectrograph (or spectrogram)” is “a representation of an utterance of speech (or roar of a tiger), laying out as a function of time (on the x-axis) the frequencies contained in the utterance (on the y-axis), and the loudness (indicated by the darkness of the tracing). In other words, it is a graphical representation of loudness and frequency as a function of time.” Thomas L. Bohan, *Scientific Evidence and Forensic Science Since Daubert: Maine Decides to Sit Out the Dance*, 56 ME. L. REV. 101, 134 (2004).

different sound, independently generated from this representation or chart is also not the “actual sound” originally “fixed” in the sound recordings. In fact, this Court already found that even *actual* audio recordings that “mimic attributes of their voices like “pitch, loudness, tone, timbre, cadence, inflection, breathiness, roughness, strain, jitter, (variation in pitch), shimmer (variation in amplitude), spectral tilt, and overall intelligibility, and other qualities” is not an “identical fixations of sound within the meaning of the Copyright Act.” *See Lovo*, 2025 WL 1902547, at *20 and n. 21. Those things are not the “components of a work that are original to the author” (*id.*) (citing *Boisson v. Banian, Ltd.*, 273 F.3d 262, 268 (2d Cir. 2001)). Thus, creating a chart or graphic representation of a recording’s various sound attributes is not covered by the Copyright Act. Thus, since all training, as alleged by Plaintiffs, begins with and is based on conversion of the sound recordings into spectrograms, there is no alleged copying and Plaintiffs’ Training Claim should be dismissed.

C. The Training Claim Should Be Dismissed Because the Alleged Copying Is Really The Creation of “Output” Which The Court Determined Does Not Infringe

The Court already dismissed the claim that the AI “outputs” or the “voice clones themselves” are infringing since “Copyright protection does not extend to this kind of imperfect mimicry” and that there is no protection for “voices *qua* voices” as copyright “does not extend to something as abstract and intangible as a ‘voice’” at *20 (quoting *Peter F. Gaito Architecture, LLC v. Simone Dev. Corp.*, 602 F.3d 57, 67 (2d Cir. 2010) and citing *In re Jackson*, 972 F.3d 25, 47 (2d Cir. 2020)). The Copyright office also explains that “A replica of their image or voice alone would not constitute copyright infringement.” “Copyright and Artificial Intelligence” at <https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-1-Digital-Replicas-Report.pdf>. Plaintiffs’ Training Claim fails for the same reason as Plaintiffs’ allegations of copying in the Training Claim involve the spectrograms, output, and “nascent clones”. These are

the same non-infringing imperfect mimicry already dismissed.

Plaintiffs allege that all their sound recordings were converted to spectrograms as the base for subsequent training (Cplt. ¶¶121-122). Then, Lovo created and used a “neural network” (*id.* ¶121) which “incorporates all the phonemes of the recordings”, in other words, the spectrogram data, and “replicates them when prompted” (*id.* ¶127). These replications based on the spectrograms (not copies of actual sounds) are then used purportedly in either a generative adversarial network process or a “likelihood training process” (*id.* ¶¶124-125). Plaintiffs further allege that the “the Kyle Snow and Sally Coleman voices fuse the knowledge that the Genny model learned in both the pretraining and training phases with the particular aspects of Mr. Lehrman and Ms. Sage’s voices exhibited in the Copyrighted Works” (*id.* ¶131), all through the spectrograms. The allegations are that Lovo utilizes the spectrograms and output created from the spectrograms (i.e., nascent clones) for comparison and training. There is no copying of actual sounds from the sound recordings. This is no different than the already dismissed output claim.

Further, as alleged, the training is focused on phonemes (smallest unit of sound) and characteristics like “pitch, pacing, accent, and tone” (*id.*) which are unprotectible “functional elements” of a recording. *Lovo*, 2025 WL 1902547, at *20; *see also Gaito*, 602 F.3d at 68 (citing *Attia v. Soc’y of New York Hosp.*, 201 F.3d 50, 55 (2d Cir. 1999)); *VMG Salsoul, LLC v. Ciccone*, 824 F.3d 871, 883 (9th Cir. 2016) (“[a] new recording that mimics the copyrighted recording is not an infringement, even if the mimicking is very well done, so long as there was no actual copying.”).

While there are some allegations about the “duplication”, “rearranging”, “recapturing”, “remixing”, and/or “reprocessing” sound recordings (Cplt. ¶¶123, 126, 133, 254-256), those (conclusory) assertions appear to be alleged to be tied to the processes of converting sound

recordings to a spectrogram and using that spectrogram to create “nascent clones” and other output based on the phonemes and other attributes from the spectrogram. This output is then allegedly used for training. The dependence of Training Claim on the dismissed output claim is made clear by the allegation in paragraph 257 that the Training Claim is based on “Defendant’s reproduction, distribution, public display, and sale of the entirety and portions of the Copyrighted Works without Plaintiffs’ permission”. This claim is functionally the same as the “output” infringement claim and has already been dismissed.

Plaintiffs’ Training Claim should be dismissed with prejudice.

D. The Training Claim Should Be Dismissed Because the Alleged Use of the Sound Recordings Was Not Substantially Similar and Was *De Minimis*

Even if the Training Claim is deemed non-conclusory, non-speculative and to have sufficiently alleged improper copying, Plaintiffs still need to adequately allege “substantial similarity” about the copying, “(i) that it was [of] protected expression in the earlier work’ and (ii) ‘that the amount that was copied is ‘more than de minimis.’” *Gottlieb Dev. LLC v. Paramount Pictures Corp.*, 590 F. Supp. 2d 625, 631 (S.D.N.Y. 2008) (quoting *Tufenkian Import/Export Ventures, Inc. v. Einstein Moomjy, Inc.*, 338 F.3d 127, 131 (2d Cir.2003)); see also *Feist*, 499 U.S. at 361 (holding that copyright infringement plaintiff must show valid copyright and copying of original elements of work); *DiTocco v. Riordan*, 815 F. Supp. 2d 655, 666 (S.D.N.Y. 2011), *aff’d*, 496 F. App’x 126 (2d Cir. 2012); *Bell v. Blaze Magazine*, 2001 WL 262718, at *3 (S.D.N.Y. Mar. 16, 2001) (dismissing claim where no reasonable jury could find substantial similarity or that similarity pertains only to unprotected aspects of work). “To determine similarity, courts ask whether ‘an average lay observer would recognize the alleged copy as having been appropriated’”. *Lovo*, 2025 WL 1902547, at *18 (quoting *Hamil Am. Inc. v. GFI*, 193 F.3d 92, 100 (2d Cir. 1999)). Here, Plaintiffs have not plausibly alleged either that any “copying” was of “protected expression”

or that what was copied is “more than *de minimis*”.

First, Plaintiffs have not plausibly alleged that there was protected expression copied. Plaintiffs have not alleged what part of the sound recordings is allegedly protectible. The Training Claim fails for this reason alone because there is no way to determine from the Complaint if a protected element was copied. Also, the base of the Training Claim is the allegation that the sound recordings were converted to spectrograms. Even if this conversion constitutes factual copying (it does not), there can be no allegation that anything protectible was copied in the conversion. The spectrogram conversion only converts measurable attributes of a voice like those already found to be not protected (i.e. “pitch, loudness, tone, timbre, cadence, inflection, breathiness, roughness, strain, jitter, (variation in pitch), shimmer (variation in amplitude), spectral tilt and overall intelligibility” *Lovo*, 2025 WL 1902547, at *20); *see also New York Times Co. v. Microsoft Corp.*, 777 F. Supp. 3d 283, 328 (S.D.N.Y. 2025) (claim dismissed because summaries “which differ in style, tone, length, and sentence structure from [originals]—are not substantially similar” to original, “qualitatively or quantitatively”). Similarly, anything created from a spectrogram is similarly limited. This is true also for “phonemes” or “nascent clones”. *See, e.g., Gottlieb*, 590 F. Supp. 2d at 633 (“average lay observer” would not discern any protected elements in copied material); *DiTocco*, 815 F. Supp. 2d at 673 (similarities “relate essentially to nonprotectible elements, and the similarities in the works' expression are greatly outweighed by their differences); *Dobson v. NBA Props., Inc.*, 1999 WL 97901, at *2 (S.D.N.Y. Feb. 24, 1999) (claim dismissed because copied material not copyrightable); *Boyle v. Stephens, Inc.*, 1998 WL 80175, at *6 (S.D.N.Y. Feb. 25, 1998) (infringement claim dismissed where “no reasonable juror” could find substantial similarity between protected expressive elements of plaintiff's and defendant's works). Thus, there is no plausible allegation that there was any copying of protected material and the

Training Claim should be dismissed.

Second, even if the Court deems that there was copying of protected material, there is no plausible allegation that there was copying of any original or protected material was more than *de minimis*. “The legal maxim ‘*de minimis non curat lex*’—‘the law does not concern itself with trifles’—applies in the copyright context ... if the copying is *de minimis* and so “trivial” as to fall below the quantitative threshold of substantial similarity, the copying is not actionable”. *Gottlieb*, 590 F. Supp. 2d at 632 (granting 12(b)(6) dismissal); *see also Gayle v. Allee*, 2021 WL 120063, at *8 (S.D.N.Y. Jan. 13, 2021) (“the use of the ‘Art We All One’ phrase was both qualitatively and quantitatively *de minimis* such that it cannot support an infringement claim”); *TufAmerica, Inc. v. WB Music Corp.*, 67 F.Supp.3d 590, 598-599 (S.D.N.Y. 2014) (“no plausible claim of substantial similarity” and even if there is “factual copying”, infringement requires “that the copying is quantitatively and qualitatively sufficient to support the legal conclusion that infringement ... has occurred”).

Here, Plaintiffs allege that there was conversion to spectrograms, which labels “sound waves and their corresponding phoneme, the smallest unit of a sound in language” (Cplt. ¶121) and that “the neural network incorporates all the phonemes of the recordings” (*id.* ¶127). These small units of sounds, even if deemed copying of protected material, are too small to allow a lay observer determine that there is substantial similarity to any protected material. *See, e.g., TufAmerica*, 67 F. Supp. 3d at 595–96 (2 minutes and 35 seconds was *de minimis*, and there was “no plausible claim of substantial similarity here”); *see also Ong v. Townsquare Media, Inc.*, 2025 WL 2294867, at *4 (S.D.N.Y. Aug. 8, 2025) (“[T]he single frame used ... constitutes less than one one-thousandth of the total number of frames in the sixty-three-second video ... [which] was *de minimis*.”) (cleaned up); *Richardson v. Townsquare Media, Inc.*, 2025 WL 89191, at *4

(S.D.N.Y. Jan. 14, 2025) (“Single still-frame screenshots from videos ... are generally considered to constitute *de minimis* use and fall below the substantial similarity threshold”); *Rudkowski v. MIC Network, Inc.*, 2018 WL 1801307, at *4 (S.D.N.Y. Mar. 23, 2018) (“use of Plaintiff’s Video is trivial. An individual frame clearly represents an extremely small fragment of the whole Video”); *Gottlieb*, 590 F. Supp. 2d at 633 (logo which “occupies less than five percent of the frame area” is *de minimis*). Thus, this type of “copying” of phonemes (via spectrograms) to train artificial intelligence and create nascent clones is trivial and *de minimis*.

The Training Claim should be dismissed with prejudice for this reason.

E. The Training Claim Should Be Dismissed Because Any Alleged Copying Is “Fair Use”

Even if the Court finds that there was potentially non-*de minimis* infringing copying, the Training Claim should still be dismissed because using the sound recordings for training is clearly a “fair use”. *See Kelley v. Morning Bee, Inc.*, 2023 WL 6276690, at *11 (S.D.N.Y. Sept. 26, 2023) (dismissing copyright claims as “fair use”); *Santos v. Kimmel*, 745 F. Supp.3d 153, 166 (S.D.N.Y. 2024) (same); *Schwartzwald v. Oath Inc.*, 2020 WL 5441291, at *9 (S.D.N.Y. Sept. 10, 2020) (same); *Yang v. Mic Network, Inc.*, 405 F. Supp. 3d 537, 540 (S.D.N.Y. 2019) (same).

The fair use doctrine is a statutory exception to copyright infringement, permitting the unauthorized use of a protected work for certain purposes. 17 U.S.C. § 107. These may include, for example, “purposes such as criticism, comment, news reporting, teaching ..., scholarship, or research.” *Id.* In this way, “[t]he fair use doctrine ‘permits courts to avoid rigid application of the copyright statute when, on occasion, it would stifle the very creativity which that law is designed to foster.’” *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 527 (2023) (citing *Stewart v. Abend*, 495 U.S. 207, 236 (1990)). “[T]he fair use determination is an open-ended and context sensitive inquiry,” in which courts weigh four non-exclusive statutorily

provided factors in light of the purposes of copyright. *Kelley*, 2023 WL 6276690, at *10 (quotation omitted).

“The statutory framework provides four nonexclusive factors courts should consider when deciding if a particular use is ‘fair’: (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes; (2) the nature of the copyrighted work; (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and (4) the effect of the use upon the potential market for or value of the copyrighted work.” *Wilder v. Hoiland*, 2025 WL 783642, at *2 (2d Cir. Mar. 12, 2025). “Fair use” is a “flexible” concept. *Id.*

The “first” factor, the purpose and character of the use clearly favors fair use for the alleged Training Claim. This factor “focuses on whether an allegedly infringing use has a further purpose or different character, which is a matter of degree, and the degree of difference must be weighed against other considerations, like commercialism”. *Warhol*, 598 U.S. at 525. “The “central” question it asks is whether the new work merely ‘supersede[s] the objects’ of the original creation ... (‘supplanting’ the original), or instead adds something new, with a further purpose or different character.” *Id.* at 528. “A use that has a further purpose or different character is said to be ‘transformative,’” and “the degree of transformation required to make ‘transformative’ use of an original must go beyond that required to qualify as a derivative.” *Id.* at 510.

Here, Plaintiffs allege that the relevant sound recordings were not intended to be used for training AI to generate artificial voices and recordings. As alleged, (a) Lehrman created sound recordings by reading and recording scripts provided to him by Lovo intending them to be used “for ‘research purposes only,’ and that such research would be only ‘internal’ and ‘academic’ in nature” (*Lovo*, 2025 WL 1902547, at *2; Cplt. ¶¶48-50); and (b) Sage created sound recordings

by reading and recording scripts provided by Lovo to be used as “test scripts for radio ads” which “will not be disclosed externally” (*Lovo*, 2025 WL 1902547, at *2; Cplt. ¶¶75-78). This Court already determined that both academic research and “test scripts for radio ads” “do[] not include the creation of commercial AI models”. *Lovo*, 2025 WL 1902547, at *19, fn. 19. It is clear that, as alleged, (a) neither Sage nor Lehrman created their sound recordings to be anything other than recordings for one-time usage; and (b) neither Lehrman nor Sage created their sound recordings to be used commercially at all. Thus, Lovo’ alleged use of the sound recordings to train its commercial AI model is “plainly different from the original purpose for which they were created.”. *Kelley*, 2023 WL 6276690, at *12 (dismissing claim because fair use not alleged) (quoting *Bill Graham Archives v. Dorling Kindersley Ltd.*, 448 F.3d 605, 609 (2d Cir. 2006)).

As the court explained in *Bartz v. Anthropic PBC*, “the purpose and character of using copyrighted works to train LLMs to generate new text was quintessentially transformative. Like any reader aspiring to be a writer, Anthropic's LLMs trained upon works not to race ahead and replicate or supplant them — but to turn a hard corner and create something different. If this training process reasonably required making copies within the LLM or otherwise, those copies were engaged in a transformative use”. 2025 WL 1741691, at *8 (N.D. Cal. June 23, 2025); *see also Kadrey v. Meta Platforms, Inc.*, 2025 WL 1752484, at *9 (N.D. Cal. June 25, 2025) (“There is no serious question that Meta's use of the plaintiffs’ books had a ‘further purpose’ and ‘different character’ than the books—that it was highly transformative. The purpose of Meta's copying was to train its LLMs, which are innovative tools that can be used to generate diverse text and perform a wide range of functions”). Here too the purpose and character of the alleged use of copyrighted sound recordings to train AI to generate new sounds is quintessentially transformative and not to replicate or supplant the original sound recordings but to create something different. Also, as in

Anthropic, there is no question here that the “output provided to users” does not infringe. *Anthropic*, 2025 WL 1741691, at *7. The Court, here, has already dismissed any output infringement claim so, like in *Anthropic*, the only issue is the input for training.

Moreover, here, there is an additional reason to find that this first factor favors fair use. Here the alleged “copying” is not just transformative in purpose, but it was actually transformative to the sound recording itself by transforming the recording into a spectrogram and into phonemes. The alleged training having converted the sound recordings into visual representations and atomized data points, then converted those data points back into (new, independently created) sound, and then from there on out, generated wholly new creations altogether. This is clearly transformative. Thus, the first factor weighs heavily in favor of Lovo’s fair use of the sound recordings.

The “second” factor is the “nature of the copyrighted work” and it “rarely play[s] a significant role” in determining “fair use”. See, e.g., *Yang*, 405 F. Supp. 3d at 546; *Clark v. Transportation Alternatives, Inc.*, 2019 WL 1448448, at *4 (S.D.N.Y. Mar. 18, 2019) (same); *Whiddon v. BuzzFeed, Inc.*, 638 F. Supp. 3d 342, 354 (S.D.N.Y. 2022) (similar). However, here it also favors fair use slightly because the courts “give ‘greater leeway ... to a claim of fair use where the work is factual or informational’ as opposed to ‘expressive or creative.’” *Wilder*, 2025 WL 783642, at *3 (quotation omitted). Here, there is no question there is little, if any, expressive or creative content to sound recordings of scripts provided by Lovo sold for \$400 or \$1200 on Fiverr to be used for internal research or test radio ads and then not used again. Thus, this sub-factor favors fair use. However, courts also consider “whether the work is published or unpublished” with “fair use involving unpublished works being considerably narrower.” *Santos*, 745 F. Supp.3d at 165-166 (quotation omitted). These sounds recordings are unpublished. But, as alleged,

Plaintiffs never intended to sell these publicly as they were only recorded for Lovo from Lovo scripts and thus differ from other unpublished more clearly creative material. Thus, the nature of sound recordings generally and these in particular makes this factor slightly favor fair use.

The “third” factor which is “the amount and substantiality of the portion” of the copyrighted work used and was it “reasonable in relation to the purpose of the copying”. *Kadrey*, 2025 WL 1752484, at *14. As argued *infra* at Part II(D), the alleged copied material is not substantially similar. But, “[w]hat matters [] is not so much ‘the amount and substantiality of the portion used’ *in making a copy*, but rather the amount and substantiality of *what is thereby made accessible* to a public for which it may serve as a competing substitute” *Authors Guild v. Google, Inc.*, 804 F.3d 202, 222 (2d Cir. 2015) (cleaned up). Here, the Court has already determined that there is no viable output claim here so none of the sound recordings will be made accessible to the public; the Training Claim is supposed to be only for training purposes. Thus, this factor also favors fair use. *See, e.g., Anthropic*, 2025 WL 1741691, at *15 (since no infringement of output, copying “to train the LLMs underlying Claude” was “especially reasonable”); *Kadrey*, 2025 WL 1752484, at *14 (third factor in training AI favors fair use because it was reasonable).

The “fourth” factor considers “the effect of the use upon the potential market for or value of the copyrighted work.” 17 U.S.C. § 107(4). The relevant question is “if the challenged use becomes widespread, it will adversely affect the potential market for the copyrighted work.” *Kelley*, 2023 WL 6276690, at *14 (quoting *Andy Warhol Found. for Visual Arts, Inc. v. Goldsmith*, 11 F.4th 26, 48 (2d Cir. 2021)). In *Anthropic*, the court determined that since there was no output infringement claim, there could be no impact on the market for the original works and “[t]he copies used to train specific LLMs did not and will not displace demand for copies of Authors’ works, or not in the way that counts under the Copyright Act.” 2025 WL 1741691, at *16. Here, too, there

is no output infringement claim. Thus, the “copies” allegedly used here for training are not released to the public and cannot compete with the original sound recordings. But, here, this factor favors fair use even more strongly because, as alleged, there was and is no market for the original sound recordings. They were a work-for-hire created specifically and only for Defendants and not for sale to anyone else. Thus, there is no market effect at all. The fourth factor, thus, strongly favors fair use.

Therefore, all four factors favor fair use on the face of the complaint. The Court should accordingly dismiss the Training Claim.

CONCLUSION

For the foregoing reasons, Defendant respectfully requests that the Court grant Defendant’s motion to dismiss the Training Claim (the Sixth Cause of Action) with prejudice.

Dated: September 15, 2025
New York, New York

RIMÔN, P.C.

By: /s/Michael S. Lazaroff
Michael S. Lazaroff
400 Madison Ave, Suite 11D
New York NY 10017
646-738-4151
michael.lazaroff@rimonlaw.com

David E. Case
800 Oak Grove Avenue, Suite 250
Menlo Park, California 94025
408-512-2239
david.case@rimonlaw.com

William W. Bergesch
100 Jericho Quadrangle, Suite 300
Jericho, New York 11753
516-479-6300
william.bergesch@rimonlaw.com

Attorneys for Defendant Lovo, Inc.

CERTIFICATE OF COMPLIANCE

I certify that this memorandum of law complies with Rule 4(C) of the Individual Practices of the Honorable J. Paul Oetken and contains **8,711** words, inclusive of footnotes and endnotes, and exclusive of the caption, table of contents, table of authorities, signature blocks, and this certificate, based on the word count tool of the word processing program used to prepare this document, and otherwise complies with Southern District of New York Local Civil Rule 7.1(c).

/s/ Michael S. Lazaroff _____
Michael S. Lazaroff