PUBLIC REDACTED VERSION

SONAL N. MEHTA (SBN 222086)
sonal.mehta@wilmerhale.com
WILMER CUTLER PICKERING HALE AND DORR LLP
2600 El Camino Real, Suite 400
Palo Alto, CA 94306
Telephone (650) 858-6000

TAYLOR GOOCH (SBN 294282)
taylor.gooch@wilmerhale.com
WILMER CUTLER PICKERING HALE AND DORR LLP
50 California Street
San Francisco, CA 94111
Telephone: (628) 235-1000

ALEX W. MILLER (*pro hac vice*)
alex.miller@wilmerhale.com
JARED V. GRUBOW (*pro hac vice*)
WILMER CUTLER PICKERING HALE AND DORR LLP
jared.grubow@wilmerhale.com
250 Greenwich Street
New York, NY 10007
Telephone: (212) 230-8800

LOUIS W. TOMPROS (*pro hac vice*)
louis.tompros@wilmerhale.com
WILMER CUTLER PICKERING HALE AND DORR LLP
60 State Street
Boston, MA 02109
Telephone: (617) 526-6000

ARI HOLTZBLATT (SBN 354361)
ari.holtzblatt@wilmerhale.com
ROBIN C. BURRELL (*pro hac vice*)
WILMER CUTLER PICKERING HALE AND DORR LLP
2100 Pennsylvania Ave, NW
Washington, DC 20037
Telephone: (202) 663-6000

Attorneys for Defendant
ANTHROPIC PBC

UNITED STATES DISTRICT COURT

NORTHERN DISTRICT OF CALIFORNIA, SAN JOSE DIVISION

CONCORD MUSIC GROUP, INC., ET AL.,

Plaintiffs,

v.

ANTHROPIC PBC,

Defendant.

Case Number: 5:24-cv-03811-EKL-SVK

ANTHROPIC'S REPLY SEPARATE STATEMENT IN SUPPORT OF MOTION FOR PARTIAL SUMMARY JUDGMENT

Judge Eumi K. Lee
Magistrate Judge Susan van Keulen

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
Anthropic's Use Of Lyrics As Inputs To Train Claude Is Fair Use		
Issue 1: The Use of Lyrics as Inputs to Training is Fair Use as Matter of Law	Fact 1: Musician and rapper, Meek Mill, uses Claude to help him understand and manage the administrative aspects of his music career. Ex. 145 (MeekMill (@MeekMill) Twitter (Mar. 23, 2026), https://x.com/MeekMill/status/2036107917095469093)[1]	Disputed. The cited evidence states only that, as of March 23, 2026, "Claude is helping [Meek Mill] organize [his] whole music career and other businesses in days. . . and it's moving [his] business forward at a high rate!"
	Fact 2: Anthropic co-founder and Chief Science Officer, Dr. Jared Kaplan, testified that Anthropic was founded to "form an organization that would focus more fully on safe and beneficial AI development." Kaplan testified that his Anthropic co-founders were "Chris Olah, Jack Clark, Sam McCandlish, Tom Brown, Dario Amodei, and Daniela [Amodei]." Ex. 120 (Kaplan Dep.) at 62:18-71:16	Undisputed that Dr. Kaplan testified as quoted.
	Fact 3: Anthropic co-founder and Chief Science Officer, Dr. Jared Kaplan, testified that "the overall view of the founders in 2021, I would say, was strongly against commercialization, and towards the idea that we would effectively never pursue it, and that we might be funded primarily by philanthropists in perpetuity." Ex. 120 (Kaplan Dep.) at 115:11-16	Undisputed that Dr. Kaplan testified as quoted.

[1] Unless otherwise specified, Anthropic's citations to 'Ex.' are exhibits attached to the Declaration of Robin C. Burrell ("Burrell Declaration") in support of Anthropic's Motion for Summary Judgment and Opposition to Publishers' Motion for Partial Summary Judgment filed concurrently herewith. Citations to 'Ant. Reply Ex.' are exhibits to the Declaration of Robin C. Burrell filed concurrently with Anthropic's Reply Brief in Support of its Motion for Summary Judgment.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	Fact 4: Anthropic co-founder and President, Ms. Daniela Amodei, testified "when we first founded the company in 2021, we were primarily focused on research at that time, and we always theorized that had we would have, you know, potential commercial applications for the models down the road." Ex. 111 (Daniela Amodei Dep.) at 120:8-12	Undisputed that Ms. Amodei testified as quoted.
	Fact 5: Anthropic co-founder, Mr. Tom Brown, testified that "[i]nitially we weren't sure if we were even going to commercialize anything, and so we had kind of a long soul-searching process." Ex. 106 (Brown Dep.) at 65:10-12	Undisputed that Mr. Brown testified as quoted.
	Fact 6: Anthropic co-founder, Mr. Benjamin Mann, testified "So one of Anthropic's principles is what we call the race to the top. Our hope is that by showing that we, as one of the frontier labs, can prioritize both safety and capabilities at the same time, that we can put pressure on our competitors to also prioritize safety." Ex. 124 (Mann 30(b)(1) Dep. Vol. 2) at 188:18-189:16	Undisputed that Mr. Mann testified as quoted. Anthropic relies on this fact for the truth of the matter asserted by the quoted testimony. Publishers dispute the truth of the matter asserted that Anthropic "prioritize[s] both safety and capabilities at the same time." Anthropic has prioritized commercial goals over research and safety goals. **Pls.' Ex. 46** (Dep. Ex. 173, Anthropic_0000443204); **Burrell Decl., Ex. 111** (Daniela Amodei Dep. Tr.) at 162:9–164:12 (discussing Dep. Ex. 173); **Pls.' Opp'n/Reply Ex. 157** (Dep. Ex. 237, Anthropic_0000479241) at -242, -245; **Pls.' Opp'n/Reply Ex. 128** (Kaplan Dep. Tr.) at 202:14–203:22 (discussing Dep. Ex. 237); **Pls.' Opp'n/Reply Ex. 156** (Dep. Ex. 235, Anthropic_0000215992); **Pls.' Opp'n/Reply Ex. 126** (Dario Amodei Dep. Tr.) at 125:1–129:7 (discussing

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
		Dep. Ex. 235); **Pls.' Opp'n/Reply Ex. 148** (Dep. Ex. 35, Anthropic 0000229152) at -152; **Pls.' Ex. 7** (Turvey 30(b)(6) Dep. Tr.) at 109:3–110:18; *see also* **Pls.' Opp'n/Reply Ex. 127** (Daniela Amodei Dep. Tr.) at 147:17–48:2 ("Q When [Dario Amodei] says 'future deployment,' that's referring to deploying commercial products, correct? A Likely, yes. Q He wants Anthropic's commercial products to [redacted] That's what he writes, correct? A He wrote [redacted] among other things in [Dep. Ex. 170]."); **Pls.' Ex. 43** (Dep. Ex. 170, Anthropic_0000458427) (July 2021 memo authored by Dario Amodei discussing whether Anthropic can "reduce data recitation/solve the copyright problem internal to the model," which "is a problem that's clearly on people's minds in both a commercial and research sense"). Publishers further dispute the truth of the matter asserted that Anthropic "prioritize[s] both safety and capabilities at the same time." In February 2026, Anthropic dropped its flagship safety commitment to responsible AI scaling, including because its founders "didn't really feel, with the rapid advance of AI, that it made sense for us to make unilateral commitments . . . if competitors are blazing ahead." **Pls.' Opp'n/Reply Ex. 178** (Billy Perrigo, *Exclusive: Anthropic Drops Flagship Safety Pledge*, TIME (Feb. 24, 2026), https://time.com/7380854/exclusive-anthropic-drops-flagship-safety-

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
		pledge/) (quoting interview with Anthropic founder Jared Kaplan); **Pls.' Opp'n/Reply Ex. 179** (ANTHROPIC, *Anthropic's Responsible Scaling Policy: Version 3.0* (Feb. 24, 2026), https://www.anthropic.com/news/responsible-scaling-policy-v3); *but see* **Pls.' Ex. 95** (Dep. Ex. 247, *Transcript: Ezra Klein Interviews Dario Amodei,* N.Y. TIMES (Apr. 12, 2024), https://www.nytimes.com/2024/04/12/podcasts/transcript-ezra-kleininterviews-dario-amodei.html) at -9–10 (Dario Amodei interview discussing importance of responsible AI scaling and commitment to pausing unsafe AI development); **Pls.' Opp'n/Reply Ex. 126** (Dario Amodei Dep. Tr.) at 45:17–46:22; **Burrell Decl., Ex. 125** (Mann 30(b)(1) Dep. Tr.) at 187:11–188:4 ("[W]e've publicly committed to potentially shutting down our models in the future for the right reasons").
	Fact 7: Anthropic co-founder and Chief Science Officer, Dr. Jared Kaplan, testified that Anthropic's status as a public benefit corporation allows it to prioritize its mission above its profit motive. Ex. 120 (Kaplan Dep.) at 106:5-24	Disputed. When asked, "What is a public benefit corporation?", Dr. Kaplan testified, "My general understanding is that a public benefit corporation is a corporation that can have a mission as its highest principle, and that it can legally, operationally, as—as an entity hold its mission as its highest goal above profit motives." But when asked, "Does Anthropic have a formal mission as its highest goal as a public benefit corporation?", Dr. Kaplan testified, "I don't recall." And when asked, "[Y]ou don't know whether Anthropic has a mission as a public benefit corporation?", Dr. Kaplan testified, "I don't recall the legal

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
		details of how that is organized or set up." **Burrell Decl., Ex. 120** (Kaplan Dep.) at 106:13–108:10.
	Fact 8: Anthropic's certificate of incorporation states "The specific public benefit that the Corporation will promote is to responsibly develop and maintain advanced AI for the cultural, social and technological improvement of humanity." Ex. 88 (Anthropic_0000570401) at -402	Undisputed that Anthropic's certificate of incorporation contains the quoted language.
	Fact 9: Anthropic co-founder and Chief Science Officer, Dr. Jared Kaplan, attested that Claude is trained to be both helpful and harmless. Dkt. 212-5 ("Kaplan PI Decl.") at ¶¶ 39-41	Undisputed that Dr. Kaplan attested that "Claude is trained to be both helpful and harmless."
	Fact 10: Publishers' expert, Dr. Shawn Shan, reported: "Text-generative LLMs like Claude are a class of generative AI that are trained on and mimic natural human language in their written outputs." And that "LLMs like Claude analyze the statistical relationship or conditional probability of tokenized words and phrases and are taught to generate written outputs based on preceding context. In the case of LLMs like Claude, the outputs are texts that appear natural and coherent, generated in response to human inputs."	Undisputed.

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	Ex. 7 (Shan Report) at ¶ 19	
	Fact 11: Publishers' expert, Dr. Ben Zhao, reported that "generative AI systems are trained on very large datasets of content according to the form of content they are intended to generate and input they are designed to accept." He also testified "Claude's training appears to be consistent with the general description above." Ex. 10 (Zhao Report) at ¶¶ 23, 55	Undisputed, except that Dr. Zhao did not "testif[y]" as quoted. The quoted language appears in Dr. Zhao's opening expert report. **Burrell Decl., Ex. 10** (Zhao Report) at ¶ 55.
	Fact 12: Publishers' expert, Dr. Ben Zhao, reported that "LLMs are autoregressive transformers that extract probabilities between words and contextual cues, in order to predict words, phrases and larger pieces of text." Ex. 18 (Zhao Rebuttal) at ¶130	Undisputed.
	Fact 13: Anthropic co-founder and Chief Science Officer, Dr. Jared Kaplan, attested that "LLMs are text-based generative AI models that use deep learning techniques and train on extremely large data sets to understand, summarize, generate, and predict new content." He also attested "Anthropic is an artificial intelligence safety and research company based in San Francisco, California, working to develop generative AI models." Claude was released to the general public in July 2023. Kaplan PI Decl. ¶¶ 6, 7, 8	Undisputed that Dr. Kaplan attested as quoted. Also undisputed that Anthropic is an artificial intelligence company based in San Francisco, CA. Anthropic relies on this fact for the truth of the matter asserted by the quoted attestation. Publishers dispute the truth of the matter asserted that LLMs "generate[] and predict new content." In the large language model ("LLM") research community, it is still uncertain whether an LLM is even capable of generating "new" content (as opposed to memorizing and regurgitating its training data). Decl. of Ben Zhao ISO Pls.' Mot. for Summ. J. (March 23, 2026), ¶¶ 44–47, ECF No. 604 ("Zhao Decl."); **Burrell Decl., Ex. 10** (Zhao Report) ¶¶ 26, 46, 91.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
		Disputed that "Claude was released to the general public in July 2023." Claude was released to the general public in March 2023, and Claude 2 was released in July 2023. Answer to FAC ¶ 54; **Pls.' Ex. 73** (Dep. Ex. 187, Anthropic_0000349548); *see also* Plaintiffs' Moving Separate Statement ISO Mot. for Summ. J. (March 23, 2026), Facts Nos. 124, 154 (undisputed), ECF No. 594-1 ("MSS").
	Fact 14: Claude has been used for various tasks by pharmaceutical researchers, scientists, and government agencies. Ex. 1 (Altman Report) at ¶¶ 54-116	Disputed only to the extent that the cited evidence (Burrell Decl., Ex. 1) is subject to a pending motion to exclude. *See* ECF No. 640.
	Fact 15: Anthropic co-founder and President, Ms. Daniela Amodei, testified that "Research and Development, Healthcare, and Life Sciences" are among "the most common use cases of Claude." Ex. 111 (Daniela Amodei Dep.) at 116:10-117:2, 154:4-15, 197:8-17	Undisputed that Ms. Amodei testified as quoted.
	Fact 16: A December 2024 report of aggregated insights across a random sample of consumer (i.e., non-business) users' Claude conversations shows that the most popular consumer use cases for Claude were "web and mobile application development assistance" (>10%),"content creation and communication assistance across disciplines" (~9%), "multidisciplinary academic research and writing assistance" (~7%), "education and career development assistance" (~6%), "implement and optimize diverse advanced AI/ML technologies and	Undisputed that the cited Anthropic report contains the quoted language.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	applications creation" (~6%), "business strategy and operations assistance across industries" (~6%), "multilingual N[atural] L[anguage] P[rocessing], translation, and linguistic analysis services" (~4.5%), "DevOps and cloud infrastructure implementation and troubleshooting" (~4%), "digital marketing and SEO optimization assistance" (~4%), and "data analysis, visualization, and management assistance" (~4%). The top ten use cases collectively account for roughly 60% of Claude conversations. Ex. 47 (Anthropic_0000012987) at -994	
	Fact 17: Claude has been used to analyze data for clinical trials and to draft clinical study reports for medical treatments. Ex. 1 (Altman Report) at ¶ 76; Ex. 87 (Anthropic_0000570252) at -253	Disputed only to the extent that the cited evidence (Burrell Decl., Ex. 1) is subject to a pending motion to exclude. *See* ECF No. 640.
	Fact 18: Claude has been applied to help compose lyrics by brainstorming interesting word pairings. Ex. 1 (Altman Report) at ¶ 111; Ex. 98 (Anthropic_0000594977) at -982	Disputed only to the extent that the cited evidence (Burrell Decl., Ex. 1) is subject to a pending motion to exclude. *See* ECF No. 640.

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	Publishers' Additional Fact #1: Anthropic is incorporated as a "public benefit corporation." **Burrell Decl., Ex. 120** (Kaplan Dep.) at 106:5–12; **Burrell Decl., Ex. 88** (Anthropic_0000570401) at -402.	Undisputed.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	Publishers' Additional Fact #2: As a public benefit corporation, Anthropic is a for-profit corporation. **Pls.' Ex. 3** (Whittle 30(b)(6) Dep. Tr.) at 32:18-20.	Undisputed that Anthropic is a public benefit corporation and that public benefit corporations are organized as for-profit corporations. Further undisputed that, as a public benefit corporation, Anthropic is "able to make sure that [Anthropic's] mission comes first. And that if [Anthropic has] to do something that would be for the mission, but not optimizing shareholder value, [Anthropic has] the ability to do that." Ex. 88 (Anthropic_0000570401) at -402; Ex. 106 (Brown Dep.) at 51:18-52:9; Ex. 120 (Kaplan Dep.) at 105:18-107:18.
	Publishers' Additional Fact #3: As a public benefit corporation, Anthropic is not a non-profit organization. **Burrell Decl., Ex. 120** (Kaplan Dep.) at 106:25–107:10; **Pls.' Opp'n/Reply Ex. 129** (Brown Dep. Tr.) at 56:7-15.	Undisputed.
	Publishers' Additional Fact #4: Anthropic is a for-profit corporation whose certificate of incorporation permits Anthropic to "engage in any lawful act or activity for which corporations may be organized under the Delaware General Corporation Law[.]" **Burrell Decl., Ex. 88** (Anthropic_0000570401) at -402; **Burrell Decl., Ex. 120** (Kaplan Dep.) at 106:25–107:10; **Pls.' Ex. 3** (Whittle 30(b)(6) Dep. Tr.) at 32:18–20.	Undisputed that Anthropic is a public benefit corporation and the certificate of incorporation includes the quoted language. Ex. 88 (Anthropic_0000570401) at -402.
	Publishers' Additional Fact #5: Anthropic's founders began planning to commercialize Anthropic's AI products as early as March 2021.	Undisputed that Anthropic's founders "considered the possibility of commercialization" beginning in or around early 2021, as Dr. Jared Kaplan testified.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	Pls.' Ex. 43 (Dep. Ex. 170, Anthropic_0000458427, "Brainstorm on What Anthropic Might Release Externally This Year"); **Pls.' Ex. 97** (Dep. Ex. 169, Anthropic_0000444967, "Anthropic Plan and Timeline [WIP]"), at -978; **Burrell Decl., Ex. 120** (Kaplan Dep.) at 114:5–115:16; **Pls.' Opp'n/Reply Ex. 127** (Daniela Amodei Dep. Tr.) at 145:14–147:19 (testifying about Pls.' Ex. 43); **Pls.' Opp'n/Reply Ex. 164** (Anthropic_0000444301, "Anthropic 2 Year Plan and Timeline" bearing last modified date metadata of March 16, 2021) at -301, -308, -312–13; **Pls.' Opp'n/Reply Ex. 165** (Anthropic_0000444332, "Guide To Budget and 2 Year Plan" bearing last modified date metadata of March 16, 2021).	Otherwise disputed that Anthropic's founders "began planning" to commercialize at that time. "[T]he overall view of the founders in 2021, … was strongly against commercialization, and towards the idea that we would effectively never pursue it, and that we might be funded primarily by philanthropists in perpetuity" and "had a strong preference for initially creating a nonprofit organization at the time" before creating a public benefit corporation. Ant. Reply Ex. 160 (Kaplan Dep.) at 114:9-115:16, 116:9-117:11. Pls.' Ex. 43 demonstrates that Anthropic did not believe it was "logistically feasible to commercialize in 2021." Pls.' Ex. 43 (Dep. Ex. 170, Anthropic_0000458427) at -427.
	Publishers' Additional Fact #6: As of May 28, 2026, Anthropic is valued at $965 billion. **Pls.' Opp'n/Reply Ex. 191** (ANTHROPIC, *Anthropic raises $65B in Series H funding at $965B post-money valuation* (May 28, 2026)).	Undisputed.
	Publishers' Additional Fact #7: Claude has been used to look up existing song lyrics. *See, e.g.*, **Pls.' Ex. 34** (Anthropic_0000429488) at Row 3790; **Pls.' Ex. 35** (Anthropic_0000429490) at Row 3383 (similar lyric retrieval requests); **Pls.' Ex 2** (comparison chart of Publishers' lyrics and	Undisputed that Pls.' Ex. 2 shows prompts to Claude requesting lyrics. Otherwise, disputed as to the truth of the matter asserted because, to the extent that "has been used" refers to third parties, Publishers have not shown that any of the prompts in Pls.' Ex. 2 were made by bona fide third parties, including after Anthropic strengthened its guardrails.

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	Claude output in response to third-party prompts, citing Anthropic Claude records) at 26–27, 31–32, 53–54, 62–3, 82–83, 85–86, 129–30, 138–39, 256–57, 263–64, 288–89, 310–11, 317–18, 324–25; Decl. of Joshua Dennis ISO Pls.' Mot. for Summ. J. ("**Dennis Decl**.") (March 22, 2026), ¶¶ 38–39, ECF No. 602.	*See* Ex. 115 (Dennis Dep.) at 242:6-10; Ant. Reply Ex. 161 (Dennis Opening Report) ¶¶ 56-57. Pls.' Ex. 34 and Ex. 35 are cluster data from Anthropic's Clio tool and concern high-level summaries of Claude user requests, and do not support Publishers' conclusions that Claude reproduced lyrics in responses to these requests. Pls.' Ex. 34 identifies 61 records within the cluster for prompts about "Generate, modify, or retrieve song lyrics." Pls.' Ex. 35 identifies 60 records within the cluster for prompts about "Help me find, analyze, or modify song lyrics." The Clio cluster data in Pls.' Ex. 34 and Pls.' Ex. 35 show de minimis usage, representing approximately 0.2% of sampled traffic. Pls.' Ex. 34 (Anthropic_0000429488) at Row 3790; Pls.' Ex. 35 (Anthropic_0000429490) at Row 3383. Lyric related uses are very rare and not typical of Claude usage. Pls.' Opp'n/Reply Ex. 130 (Ganguli Dep.) at 59:5–11, 61:10–14; 111:19-112:16. Anthropic employs a multi-tiered defense system designed to prevent the inclusion of copyrighted material in Claude's outputs. After Publishers filed this lawsuit, Anthropic strengthened its guardrails both generally and specifically to address

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		the works in suit and stipulated with Publishers to maintain their efficacy. Dr. Bilmes's analysis shows the effectiveness of Anthropic's guardrails and defenses in preventing outputs containing the claimed lyrics to works in suit when users request existing song lyrics, as well as other purportedly copyrighted lyrics, even when directly prompted. Ex. 30 (Def.s Am. Supp. Resp. & Obj. to 1st Set of Interrogs. (Nos. 9, 11, 12) (Nov. 12, 2025)) at No. 12; Ex. 2 (Bilmes Report) at §§ V.C, IX.C; Ex. 11 (Bilmes Rebuttal) at § VI.B; Dkt. 291.
	Publishers' Additional Fact #8: Claude has been used to write purportedly "new" AI-generated song lyrics. **Pls.' Ex. 30** (D. Ganguli 30(b)(6) Dep. Ex. 6, Anthropic_0000429489) at Rows 2620, 3044, 4271, 4474, 4758; *see also* **Pls.' Ex. 31** (Anthropic_0000429482) at Rows 840, 1382, 2759, 4627 (showing clusters for lyrics-related uses); **Pls.' Ex. 32** (Anthropic_0000429485) at Rows 221, 2039, 2731, 3749, 4765 (same); **Pls.' Ex. 33** (Anthropic_0000429486) at Row 2861; **Pls.' Ex. 106** (Anthropic_0000429484) at Rows 1102, 2266, 4639 (same); Dennis Decl. ¶¶ 40, 42.	Undisputed that Claude can be used to write new and original lyrics and that Publishers' cited Clio data reflect requests for original lyrics. Otherwise disputed to the extent that Publishers imply any output contained existing lyrics. Pls.' Ex. 30, Ex. 31, Ex. 32, Ex. 33, and Ex. 106 are cluster data from Anthropic's Clio tool and concern high-level summaries of Claude user requests, and do not support Publishers' conclusions that Claude generated "new" lyrics in responses to these requests. Pls.' Ex. 30 identifies the number of records within the clusters related to prompts about "Write original song lyrics for specific themes and occasions," "Create or modify song lyrics with specific themes and requirements," "Write rock or metal song lyrics with specific themes," "Write or improve rap lyrics and songs," and "Create or improve poems and song lyrics with specific themes."

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		Pls.' Ex. 31 identifies the number of records within the clusters related to prompts about "Write original song lyrics with specific themes and styles," "Create or improve rap lyrics and compositions," "Modify or create song lyrics with specific parameters," and "Create original poems or song lyrics with specific requirements." Pls.' Ex. 32 identifies the number of records within the clusters related to prompts about "Help me create or improve song lyrics," "Write a complete song with lyrics for a specific purpose," "Write original song lyrics with specific emotional themes," "Write rap or hip-hop lyrics with specific stylistic elements," and "Create original poetry or song lyrics with specific requirements." Pls.' Ex. 33 identifies the number of records within the cluster related to prompts about "Help me find, edit, or provide song lyrics." Pls.' Ex. 106 identifies the number of records within the clusters related to prompts about "Write original song lyrics about a specific theme or person," "Write or modify song lyrics with specific themes and formats," and "Create original poems or song lyrics with specific themes." The Clio cluster data in Pls.' Ex. 30-33 and 106 show de minimis usage, representing approximately 0.2% of sampled traffic. Lyric related uses are very rare and not typical of Claude usage.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		Pls.' Opp'n/Reply Ex. 130 (Ganguli Dep.) at 59:5–11, 61:10–14; 111:19-112:16. Pls.' Ex. 30 (Ganguli 30(b)(6) Dep. Ex. 6, Anthropic_0000429489) at Rows 2620, 3044, 4271, 4474, 4758; *see also* Pls.' Ex. 31 (Anthropic_0000429482) at Rows 840, 1382, 2759, 4627; Pls.' Ex. 32 (Anthropic_0000429485) at Rows 221, 2039, 2731, 3749, 4765; Pls.' Ex. 33 (Anthropic_0000429486) at Row 2861; Pls.' Ex. 106 (Anthropic_0000429484) at Rows 1102, 2266, 4639.
	Publishers' Additional Fact #9: A "major methodological concern" with LLMs "pretrained on a broad swath of internet data, particularly large models with the capacity to memorize vast amounts of content, is potential contamination of downstream tasks by having their test or development sets inadvertently seen during pre-training." **Pls.' Opp'n/Reply Ex. 172** (Dep. Ex. 202, Tom B. Brown et al., *Language Models Are Few-Shot Learners*) at 9; *see* **Pls.' Opp'n/Reply Ex. 125** (Mann 30(b)(1) Dep. Tr.) at 26:15–27:17.	Undisputed that Pls.' Opp'n/Reply Ex. 172 contains the quoted language referring to "major methodological concern[s]" with the development of certain LLMs. Otherwise disputed as to the truth of the matter asserted as the evidence cited does not pertain to Claude. Pls.' Opp'n/Reply Ex. 172, describes training for GPT-2 and GPT-3 models (not operated by Anthropic and not at issue in this litigation), and concerns benchmark test-set contamination, not reproduction of copyrighted text in outputs, and further describes how the researchers "reduce such contamination" and discusses their "future work" to "aggressively remove data contamination." Pls.' Opp'n/Reply Ex. 172 (Dep. Ex. 202, Tom B. Brown et al., *Language Models Are Few-Shot Learners*) at 9. Mr. Mann testified that it was "correct" that Pls.' Opp'n/Reply Ex. 172 "is a research paper [he] worked

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		on while [he] [was] employed at OpenAI." Pls.' Opp'n/Reply Ex. 125 (Mann 30(b)(1) Dep.) at 26:15–27:17. As Anthropic CEO Dr. Dario Amodei and others testified, it is a "misunderstanding" that models are "memorizing or spitting things out" and Anthropic takes steps to mitigate memorization of training data and regurgitate it in outputs. Memorization is undesirable and does not "serve[] important functions" because it hurts model performance. Ex. 112 (Dario Amodei Dep.) at 94:22-25; Ex. 116 (Drain Dep.) at 66:9-23; *id.* at 83:14-18; Ex. 2 (Bilmes Report) ¶¶ 117-120; Ex. 126 (Mann 30(b)(6) Dep.) at 178:9-14; Ex. 106 (Brown Dep.) at 131:8-17.
	Publishers' Additional Fact #10: When prompted, Claude can generate output that "probably look like song lyrics . . . that look pretty good[.]" **Pls.' Ex. 4** (Mann 30(b)(6) Dep. Tr.) at 176:9-17; *see also* **Pls.' Ex. 10** (Daniela Amodei Dep. Tr.) at 180:4-6,186:10-16.	Undisputed that Mr. Benjamin Mann testified that "[i]f you ask Claude to generate song lyrics, it's likely Claude will output something that probably looks like song lyrics," and that some "instances" of these outputs "look pretty good." Further undisputed that Ms. Daniela Amodei testified that Claude is "capable of, for example, creating its own song lyrics, if it wanted to." Pls.' Ex. 4 (Mann 30(b)(6) Dep.) at 176:9-17; *see also* Pls.' Ex. 10 (Daniela Amodei Dep.) at 180:4-6, 186:10-16.
	Publishers' Additional Fact #11: Anthropic's Economic Index, which tracks occupations impacted by AI, includes data indicating that "Poets, Lyricists, and Creative	Undisputed that Anthropic's Economic Index includes data about "Poets, Lyricists, *and* Creative Writers" who have used Claude, including, as a subset of those uses, to

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	Writers" have directed Claude to "write narrative, dramatic, lyrics, or other types of poetry" and "write words to fit musical compositions, including lyrics for operas, musical plays, and choral works." **Burrell Decl., Ex. 8** (Smith Report) at ¶¶ 80, 202 & n.138.	"write narrative, dramatic, lyrics, or other types of poetry" and "write words to fit musical compositions, including lyrics for operas, musical plays, and choral works." It is also undisputed that Anthropic's Economic Index indicates that the amount of Claude usage related to the broader category of "Poets, Lyricists, *and* Creative Writers," generally, has been below 1.49%. That category expands beyond just songwriters, such that the percentage of uses related to lyrics specifically are smaller than 1.49% of uses. Ant. Reply Ex. 162 (Sept. 2025 Anthropic Economic Index, Anthropic_0000441108 at -112, -124); Ex. 8 (Smith Report) at ¶ 80, n.77 (citing Anthropic's Economic Index); Ant. Reply Ex. 163 (Smith Dep. Vol. 1) at 143:1-11.
	Publishers' Additional Fact #12: LLMs are prone to memorization, and "[l]arge LMs memorize A LOT, like a LOT." **Pls.' Opp'n/Reply Ex. 153** (Dep. Ex. 140, Anthropic_0000507750) at -796; **Pls.' Opp'n/Reply Ex. 136** (Drain Dep. Tr.) at 152:10–153:2 (testifying about Dep. Ex. 140 and explaining that "before we invested a lot more in improving our deduplication algorithms, [memorization] was a bigger problem."); **Pls.' Opp'n/Reply Ex. 135** (Sangani Dep. Tr.) at 172:5-12 ("I would not frame regurgitation as a risk. I think the risk is memorization. And if a large contiguous block of tokens has been memorized, it's—it's possible that for some combination of	Undisputed that the cited document contained the quoted language about "[l]arge LMs." Otherwise, disputed as to the truth of the matter asserted. Anthropic takes substantial steps to mitigate memorization, including through deduplicating its training data, because memorization hurts model performance. The cited observation from an Anthropic employee about LLM memorization was made in the context of deduplication research and based off a pre-trained, intermediate model, not the later commercially released model of Claude. Anthropic then used that research to investigate and to further develop techniques to reduce

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	prompts, those tokens will be reproduced. Maybe they would not be regurgitated; maybe they would. It's hard to say, but it's not unreasonable that that could occur"); *see also* Answer to FAC ¶¶ 108-09; **Pls.' Opp'n/Reply Ex. 129** (Brown Dep. Tr.) at 196:1-4 ("I think that if you trained a ton on—like, in pretraining some verbatim chunk of text many, many times, that that would make it more likely that the model would regurgitate it.").	memorization for commercially released models. Ex. 73 (Anthropic_0000507750); Ex. 116 (Drain Dep.) at 66:9-23; *id.* at 154:5-15, 178:17-179:4; Ex. 2 (Bilmes Report) ¶¶ 117-120. Further disputed as the cited testimony from Mr. Pranay Sangani and Mr. Tom Brown concerned memorization of portions of the training data as a *hypothetical* trait, not an "inherent" one. In the cited testimony, Mr. Pranay Sangani was responding to the question, "[I]f a model is engaged in memorization, does that increase the risk of potential regurgitation?" Pls.' Opp'n/Reply Ex. 135 (Sangani Dep.) at 172:1-12. Mr. Tom Brown was responding to the question, "And is it your testimony that you don't -- that you can't say one way or the other if specific content is repeated many times in the training data, whether that's more likely to influence the model's output? Mr. Brown's full response was, "I think I don't -- I think I don't know. I think that if you trained a ton on -- like, in pretraining some verbatim chunk of text many, many times, that that would make it more likely that the model would regurgitate it. But again, I think that's – that's the reason why we avoid doing stuff like that, but again, I'm not the expert here. Ant. Reply Ex. 165 (Brown Dep.) at 195:19-197:7.
	Publishers' Additional Fact #13: With respect to AI models' memorizing a particular sequence,	Undisputed that Ms. Drain testified as quoted.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	"it's good for the model to have memorized things like famous quotes that [] you or I might memorize." **Pls.' Ex. 11** (Drain Dep. Tr.) at 68:8-19.	Disputed as to the truth of the matter asserted as Anthropic takes substantial steps to mitigate memorization, including through deduplicating its training data, because memorization hurts model performance. *See* Pls.' Opp'n/Reply Ex. 136 (Drain Dep.) at 70:13-18 ("[W]e're trying to reduce memorization … because it's harmful for the model [and] it makes the model dumber[.] I don't know of any text that we've intentionally had the model memorize."); Pls.' Opp'n/Reply Ex. 135 (Sangani Dep.) at 279:24-280:8 ("We want Claude … to learn about the underlying dynamics and build an understanding for the universe that allows the model to be useful in different contexts … Memorizing data is kind of antithetical to that and is [] definitely not what we are looking for here.").
	Publishers' Additional Fact #14: Anthropic employees believe that certain song lyrics, such as those for "American Pie" by Don McLean or "Let It Be" by the Beatles, are "pillars of human culture" that reflect "useful" "cultural knowledge" for Claude to memorize verbatim. **Pls.' Opp'n/Reply Ex. 136** (Drain Dep. Tr.) at 107:24–109:10; *see also* **Pls.' Opp'n/Reply Ex. 128** (Kaplan Dep. Tr.) at 152:20–153:7.	Disputed. The deposition testimony cited does not reflect that Ms. Dawn Drain or Dr. Jared Kaplan believed that American Pie or Let It Be were "useful" "cultural knowledge" for Claude to memorize verbatim. The testimony was given only in the context of follow up questions as to their opinions as to whether songs we're "pillars of ... culture," not whether Anthropic wanted Claude to memorize them. Ms. Drain did not testify that "American Pie" or "Let It Be" were works that Anthropic wanted Claude to memorize. Rather, Ms. Drain testified, "Um, I mean, it's a pretty short list since, in general, verbatim memorization is bad and makes them all dumber, um, but, you know, one – one example we had earlier was like

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		famous bible passages that people have memorized. Um, but it would be a pretty short list" in response to the question "What are some other pillars of human culture that you would want Claude to memorize verbatim?" Pls.' Opp'n/Reply Ex. 136 (Drain Dep.) at 107:24-108:5. In response to the question, "is it true or not true that 'song lyrics may be useful to include in the training corpus because they add to the diversity of training data,'" Dr. Jared Kaplan testified, "As mentioned, I think that to understand the cultural knowledge and the way that it's situated, something like music reviews might be more useful in bridging connections between different ideas and culture. But song lyrics are also partially a source of that information, as are, for example, poems, which are very similar." Dr. Kaplan's testimony is in the context of training-data diversity, not memorization. Pls.' Opp'n/Reply Ex. 128 (Kaplan Dep.) at 152:25-153:7. Anthropic employs a multi-tiered defense system designed to prevent the inclusion of copyrighted material in Claude's outputs. After Publishers filed this lawsuit, Anthropic strengthened its guardrails both generally and specifically to address the works in suit and stipulated with Publishers to maintain their efficacy. Dr. Bilmes's analysis shows the effectiveness of Anthropic's guardrails and defenses in preventing outputs containing the claimed lyrics to works in suit, as well as other purportedly copyrighted lyrics, even when directly prompted.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		Ex. 30 (Def's Am. Supp. Resp. & Obj. to 1st Set of Interrogs. (Nos. 9, 11, 12) (Nov. 12, 2025)) at No. 12; Ex. 2 (Bilmes Report) at §§ V.C, IX.C; Ex. 11 (Bilmes Rebuttal) at § VI.B; Dkt. 291.
	Publishers' Additional Fact #15: An Anthropic employee acknowledged in an internal Claude record that a "legitimate use" for AI memorization is "memorizing longer texts that exist as one unit (lyrics, speeches, etc)," and Claude generated output in response stating that "[m]emorization serves important functions in both biological and artificial systems," including "[e]ncoding information verbatim, like lyrics or speeches, where precise recall is necessary." **Pls.' Ex. 76** (Anthropic_0000434712) at -712.	Undisputed that the cited document is a chat between an Anthropic employee and Claude, which contains the quoted language. Otherwise disputed because the cited conversation does not establish that Anthropic believed memorization was a "legitimate use" for AI. This conversation between Claude and an Anthropic employee was about the "science" of memorization. The employee also states, "yes, in neural nets, we tend to see memorization as a negative – the opposite of generalization." In response, Claude states that "excessive memorization in neural networks can hinder generalization." Pls.' Ex. 76 (Anthropic_0000434712) at -712. As Anthropic CEO Dr. Dario Amodei and others testified, it is a "misunderstanding" that models are "memorizing or spitting things out" and Anthropic does not take steps to encourage memorization in Claude; rather, it takes steps to avoid it. Memorization is undesirable and does not "serve[] important functions" because it hurts model performance. An output from Claude is not competent evidence as to LLM functionality or the purportedly "legitimate use" of memorization by LLMs.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		Ex. 112 (Dario Amodei Dep.) at 94:22-25; Ex. 116 (Drain Dep.) at 66:9-23; *id*. at 83:14-18; Ex. 2 (Bilmes Report) ¶¶ 117-120; Ex. 126 (Mann 30(b)(6) Dep.) at 178:9-14; Ex. 118 (Ganguli 30(b)(1) Dep.) at 101:5-18, 102:7-22; Ex. 106 (Brown Dep.) at 131:8-17.
	Publishers' Additional Fact #16: Songwriters have used Claude to write lyrics for songs that are then uploaded to streaming platforms. Decl. of Michael Smith ISO Pls.' Mot. for Summ. J. ("Smith Decl.") (March 23, 2026), ¶ 39, ECF No. 601.	Disputed that multiple songwriters have used Claude to write lyrics for songs that are then uploaded to streaming platforms. The cited declaration identified only one "example, in a video from June 2025" of "YouTuber Rick Beato" "us[ing] Anthropic's Claude to generate the lyrics he uses in an AI-generated song called 'West Texas Dreams,' that is currently available on … YouTube." The declaration cites no other examples and there is no evidence that "West Texas Dreams" contains the lyrics to any of Publishers' asserted works. Publishers' expert, Dr. Michael Smith testified that "You're asking an impossible question, and I don't know that I can give you a – a reasonable answer," in response to the question "[C]an you identify any songwriter or song that was displaced by Mr. Beato uploading West Texas Dreams to YouTube[?]." Publishers' expert, Dr. Michael Smith further testified, "I don't know that anybody has that – that number," in response to questioning about whether he had conducted an empirical analysis or quantification of how and whether use of Claude to write new and original song lyrics has displaced lyricists.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		Smith Decl., ¶ 39, Dkt. 601; Ant. Reply Ex. 163 (Smith Dep. Vol. 1) at 146:6-151:1; Ex. 132 (Smith Dep. Vol. 1) at 163:1-164:2.
	Publishers' Additional Fact #17: Anthropic does not "consider copyright in any of [its] helpful, honest, or harmless[ness]" trainings. **Pls.' Opp'n/Reply Ex. 125** (Mann 30(b)(1) Tr.) at 101:1–103:4; *see also* **Pls.' Ex. 74** (Dep. Ex. 179, Anthropic_0000458491) (benchmarking analysis including prompt for "awesome metal lyrics" and output containing purported lyrics on Tab "Harmlessness"); **Pls.' Ex. 4** (Mann 30(b)(6) Dep. Tr.) at 126:5-10, 127:4-7 (discussing Dep. Ex. 179).	Undisputed that Mr. Mann testified as quoted. Disputed to the extent that Anthropic does not consider copyright in training Claude. The evidence cited shows that Anthropic conducted a harmlessness assessment, in which it compared Claude's response to many prompts to test whether Claude was providing "harmful" answers using that team's specific definition of "harmfulness." Many of the prompts involved testing whether Claude or other LLMs would provide outputs that would incite violence or teach a person how to make or use weapons. One prompt in that spreadsheet of dozens of harmlessness prompts referenced "awesome Metal lyrics" to benchmark Claude's, ChatGPT's, and InstructGPT's responses. There is no evidence that these prompts were used to train Claude or that the purported lyrics in Claude's output are accurate or are a verbatim reproduction of complete lyrics to any song. Mr. Mann further testified that the "metal lyrics "look like partial quotes ... I also would be surprised if these lyrics were at all complete." Ex. 67 (Anthropic_0000458491); Ant. Reply Ex. 173 (Mann 30(b)(6) Dep.) at 130:8-133:4. In the cited testimony, Mr. Mann testified, "I don't think we would consider copyright in any of helpful,

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		honest, or harmless – this [framework] is instead more of like a baseline of what we want our models to do or not do … [Copyright is] more about our acceptable use policy and enforcing that." Pls.' Opp'n/Reply Ex. 125 (Mann 30(b)(1) Vol. 1) at 102:24-103:6. Regarding the cited benchmarking analysis, Mr. Mann testified that "Anthropic has a very specific definition of 'harmfulness' in the context of this spreadsheet, and [] this evaluation is for that definition of harmfulness.' Whether we have some opinion on copyright would be independent of this evaluation." He also testified: "[W]hen Anthropic says 'harmlessness,' we mean it in a very particular way. In our case, it's usually relating to things like hate speech, discrimination, harm to people's lives, health, emotional or physical safety, or potential for harm to other people's lives, health [or] safety, and it's certainly possible that these outputs are problematic for reasons other than harmlessness. But I feel confident that this sheet, in its rating of these things, is only considering the harmlessness aspect." Ant. Reply Ex. 173 (Mann 30(b)(6) Dep.) at 132:4-133:4 (discussing Pls.' Ex. 74 (Dep. Ex. 179, Anthropic_0000458491)). Anthropic employs a multi-tiered defense system designed to prevent the inclusion of copyrighted material in Claude's outputs. After Publishers filed this lawsuit, Anthropic strengthened its guardrails both generally and specifically to address the works in suit and stipulated with

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		Publishers to maintain their efficacy. Dr. Bilmes's analysis shows the effectiveness of Anthropic's guardrails and defenses in preventing outputs containing the claimed lyrics to works in suit when users request existing song lyrics, as well as other purportedly copyrighted lyrics, even when directly prompted. Ex. 30 (Def.s Am. Supp. Resp. & Obj. to 1st Set of Interrogs. (Nos. 9, 11, 12) (Nov. 12, 2025)) at No. 12; Ex. 2 (Bilmes Report) at §§ V.C, IX.C; Ex. 11 (Bilmes Rebuttal) at § VI.B; Dkt. 291.
	Publishers' Additional Fact #18: Anthropic's "harmlessness" training of Claude does not always prevent harmful outputs. **Pls.' Ex. 41** (Dep. Ex. 246, Anthropic_0000605700) at -740 (Slack chat from Feb. 29, 2024, in which Dario Amodei writes, [redacted] *see generally* **Pls.' Ex. 1** (comparison chart of Publishers' lyrics and Claude output in response to prompts submitted by Publishers' investigator for all 499 Works in Suit); **Pls.' Ex. 2**.	Disputed that the evidence cited shows a failure to prevent 'harmful' outputs and further disputed because the purpose of reinforcement learning with human feedback is to align Claude to human preferences of helpfulness and harmlessness not to train Claude to give specific outputs or perform specific tasks. Individuals involved in aspects of these trainings are "*instructed* to 'adversarially probe' the models to 'provoke harmful responses'" as part of the human-feedback dataset ("HF Dataset"). Anthropic's expert, Dr. Jeffery Bilmes explains that, "[t]he HF Dataset was designed to reflect human judgement of helpfulness or harmlessness across a wide range of tasks to align a model on those general human preferences. It was not built or used to train models to correctly perform any specific tasks (such as lyric retrieval)." Ex. 11 (Bilmes Rebuttal Report) ¶¶ 99, 102 (emphasis added).

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		Further disputed that the evidence cited contains "harmful" outputs rather than outputs elicited by Publishers' investigators. Ex. 20 (Bray Reply) ¶ 11. Undisputed that Anthropic will not block a launch over trivial issues. Dr. Dario Amodei testified in connection with Ex. 41: "Yes. It was -- it was my comment that -- many -- many issues come up. [redacted] Of course it is the case that we would not block a launch over tiny issues or over issues that don't have -- over issues that don't have significance. Imagine trying to run the company if, you know, we said: Oh, you know, the model shouldn't say the word 'hell' or something like that. And if the model once said the word 'hell' after we outputted a thousand pages of content from the model, then, you know, that would -- that would lead us to block the launch. That would be impractical and would – you know, the cost versus the benefit of blocking a launch under those conditions would be very lopsided." Ant. Reply Ex. 164 (Dario Amodei Dep.) at 23:4-24:2.
	Publishers' Additional Fact #19: Anthropic's President, Ms. Daniela Amodei, acknowledged that prompts relating to songs or lyrics	Undisputed that Ms. Daniela Amodei testified, "I don't believe anything related to this topic is a 'top use case.' But obviously, as you can see from

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	are "helpful" and "typical use case[s]" of Claude. **Pls.' Opp'n/Reply Ex. 127** (Daniela Amodei Dep.) at 196:16–197:21.	my using it, it's also a helpful use case to learn new artists, to get song recommendations, or to create satirical things like 'oink, oink, pink pig.'" Undisputed that, in connection with the above testimony, Ms. Amodei testified, "I believe these are typical use case[s], but probably not the most common use cases." Otherwise, disputed as to the truth of the matter asserted as Ms. Amodei was not referring to prompts relating to the song lyrics to the works in suit or song lyrics to other existing musical compositions. Ms. Amodei's testimony was made in connection with prompts to Claude to generate new and original lyrics to the tune of "Baa Baa Black Sheep." Pls.' Opp'n/Reply Ex. 127 (Daniela Amodei Dep.) at 196:16–197:21. Further disputed as to the truth of matter asserted that requests for songs or lyrics are "typical use case[s]" of Claude. Anthropic's Economic Index reveals that uses relating specifically to lyrics comprise far less than 1.49% of use cases. That category expands far beyond just songwriters, such that the percentage of uses related specifically to lyrics are smaller than at most 1.49% of uses. Ant. Reply Ex. 162 (Sept. 2025 Anthropic Economic Index, Anthropic_0000441108 at -112, -124); Ex. 8 (Smith Report) at ¶ 80, n.77 (citing Anthropic's Economic Index) Ant. Reply Ex. 163 (Smith Dep. Vol. 1) at 143:1-11.
	Publishers' Additional Fact #20: Anthropic's platform for analyzing real-world use of Claude, Clio,	Undisputed.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	which Anthropic relied on to generate the December 2024 report referenced in its Fact #16, generates only a high-level summary of "general patterns and trends" in user conversations with Claude. **Pls.' Opp'n/Reply Ex. 130** (Ganguli 30(b)(6) Dep. Tr.) at 22:19-25.	
	Publishers' Additional Fact #21: Usage insights provided by Clio are generated by Claude and do not involve human review of individual prompts and output. **Pls.' Opp'n/Reply Ex. 171** (Ganguli 30(b)(6) Dep. Ex. 5, Anthropic_0000012987) at -987, -990–91; **Burrell Decl., Ex. 119** (Ganguli 30(b)(6) Dep. Tr.) at 61:20–62:10; **Pls.' Opp'n/Reply Ex. 131** (Ganguli 30(b)(1) Dep. Tr.) at 178:15–180:19.	Undisputed.
	Publishers' Additional Fact #22: Clio is not designed to indicate whether Anthropic's guardrails are working. **Pls.' Opp'n/Reply Ex. 130** (Ganguli 30(b)(6) Dep. Tr.) at 60:6-10.	Undisputed.
	Publishers' Additional Fact #23: Clio only summarizes and groups conversations into "clusters," but does not reflect specific prompts or completions, and is in fact "designed to obfuscate" specific prompts and completions. **Pls.' Opp'n/Reply Ex. 130** (Ganguli 30(b)(6) Dep. Tr.) at 61:22–62:16, 73:1-7, 110:16-19.	Undisputed that Clio summarizes and groups conversations into clusters. Further undisputed that Dr. Deep Ganguli testified that Clio is "designed to obfuscate" specific prompts and completions as Clio was designed to respect the privacy of Claude users. Anthropic witness, Dr. Deep Ganguli testified, "I know that my team does not look at prompts and completions, especially in the context of Clio runs,

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		because Clio is designed to obfuscate that for our purposes." Dr. Deep Ganguli also explained: "Clio is not designed to look for a very specific granular detailed thing like the thing you mentioned. It's designed to analyze a large swath of usage data and find high-level patterns. And then we analyze what those high-level patterns are and -- but it's not designed to, like, answer a very specific question from usage data, if that makes sense." Dr. Deep Ganguli further stated: "we sort of designed Claude to be very privacy preserving. And one of the ways we get there is by not knowing anything about the individual conversations, just the high-level patterns and aggregate information about any cluster of conversations." Pls.' Opp'n/Reply Ex. 130 (Ganguli 30(b)(6) Dep.) at 58:15-23, 73:2-7, 110:16-19.
	Publishers' Additional Fact #24: Clio cannot conclusively prove the absence of certain patterns or conversations with Claude. **Pls.' Opp'n/Reply Ex. 130** (Ganguli 30(b)(6) Dep. Tr.) at 114:9–116:14, 125:4-18; **Pls.' Opp'n/Reply Ex. 171** (Ganguli 30(b)(6) Dep. Ex. 5, Anthropic_0000012987) at -2999.	Undisputed that the cited article states, "Furthermore, Clio cannot conclusively prove the *absence* of certain patterns or behaviors. The imperfections in the pipeline or potential adversarial actions could allow harmful behavior to escape undetected by Clio. This underscores the importance of complementing Clio's analysis with other methods." Pls.' Opp'n/Reply Ex. 171 (Ganguli 30(b)(6) Dep. Ex. 5, Anthropic_0000012987) at -2999. Undisputed that Dr. Deep Ganguli testified, "Clio is just: We take a sample of traffic. And then we sort of summarize it, and then we look for what we find. And if something is not

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		in Clio, we just don't know whether -- we just don't know much about whatever that thing that's absent is. It's unlikely that that's occurring frequently enough to show -- like, if it were occurring very frequently, it would sort of show up. But there -- you know, if it's very infrequent or if it's not happening, it won't show up." Dr. Ganguli further testified that "if something is happening and it's common … it will pop up as a cluster. If something is not happening, that's either because it's not happening or because it's just really, really infrequent." Pls.' Opp'n/Reply Ex. 130 (Ganguli 30(b)(6) Dep.) at 115:8-19.
	Publishers' Additional Fact #25: Third parties used Claude to request lyrics based upon provided song titles or lyric excerpts. **Burrell Decl., Ex. 21** (Dennis Reply) at ¶ 14.	Undisputed that Claude may be used to generate new and original lyrics based on user prompts. Mr. Dennis identified 136 records in the 5-million-record sample in which the prompt requested lyrics based on a title or lyrics. Otherwise disputed. Publishers have not shown that any bona fide "third parties" prompted Claude for "lyrics based upon provided song titles or lyric excerpts." *See* Ex. 115 (Dennis Dep.) at 242:6-10; Ant. Reply Ex. 161 (Dennis Opening Report) ¶¶ 56-57.
	Publishers' Additional Fact #26: Based on finding 136 examples of third parties requesting song lyrics based upon provided song titles or lyric excerpts in the 5-million-record Claude sample produced by Anthropic, Publishers' expert Mr. Joshua Dennis concluded that users prompted Claude for lyrics in this manner in at least [redacted] instances	Undisputed that Claude may be used to generate new and original lyrics based on user prompts. Otherwise disputed. Publishers have not shown that any bona fide "third parties" prompted Claude for "song lyrics based upon provided song titles or lyric excerpts."

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	from September 22, 2023, to March 22, 2024. **Burrell Decl., Ex. 21** (Dennis Reply) at ¶ 14.	*See* Ex. 115 (Dennis Dep.) at 241:6-243:22; Ant. Reply Ex. 161 (Dennis Opening Report) ¶¶ 56-57. Further disputed because Mr. Dennis admitted that neither the "Work-in-Suit Records" nor the "Sample Song Records" he analyzed can be extrapolated to the full set of Claude records, because neither subsample was statistically representative. And Mr. Dennis admitted that he did not verify whether those records contained Publishers' and agents' own prompts and outputs. Ex. 115 (Dennis Dep.) 79:19-25, 242:6-10, 241:6-243:22; Ex. 20 (Bray Reply) ¶¶ 47-54.
	Publishers' Additional Fact #27: Third parties have asked Claude to generate new lyrics based on or in the style of existing artists, songs, or lyrics. **Burrell Decl., Ex. 21** (Dennis Reply) at ¶ 14.	Undisputed that Claude may be used to generate new and original lyrics based on user prompts. Otherwise disputed. Publishers have not shown that that any bona fide "third parties" prompted Claude to "generate new lyrics based on or in the style of existing artists, songs, or lyrics." *See* Ex. 115 (Dennis Dep.) at 242:6-10; Ant. Reply Ex. 161 (Dennis Opening Report) ¶ 56-57.
	Publishers' Additional Fact #28: Based on finding over 600 examples of third parties asking Claude to generate new lyrics based on or in the style of existing artists, songs, or lyrics in the 5-million-record Claude sample produced by Anthropic, Mr. Dennis concluded that users prompted Claude in this manner in over [redacted] instances from September 22, 2023, to March 22, 2024.	Undisputed that Claude may be used to generate new and original lyrics based on user prompts. Otherwise disputed. Publishers have not shown that that any bona fide "third parties" prompted Claude "asking Claude to generate new lyrics based on or in the style of existing artists, songs, or lyrics." *See* Ex. 115 (Dennis Dep.) at 241:6-243:22; Ant. Reply Ex. 161 (Dennis Opening Report) ¶ 56-57.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	Burrell Decl., Ex. 21 (Dennis Reply) at ¶ 14.	Further disputed because Mr. Dennis's analysis was methodologically flawed and non-representative. Mr. Dennis's analysis is misleading in the way he categorizes users' prompts as requesting "new" lyrics. Ex. 12 (Bray Rebuttal) ¶¶13, 28; Ex. 15 (Nagaraj Rebuttal) ¶ 44.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
Issue 2: Using Lyrics to Train LLMs is Transformative	Fact 19: Publishers' expert, Dr. Ben Zhao, reported: "[G]enerative AI strives to generate content by combining existing materials together using patterns it derives from training materials. It utilizes complex machine learning architectures to extract correlation patterns between a textual prompt or request, and specific content samples (which can be text, images, or audio) that match those prompts. These models extract from training data the likelihood of a particular sequence of tokens (response) that follow other sequences of tokens (user prompt), captured and represented in layers and layers of conditional probabilities. Sufficient training on large datasets (including much of all human writing that has ever been written), produces the generative model, the encapsulation of a massive compilation of correlations between text prompts and content, stored as layers and layers of conditional probabilities." Ex. 10 (Zhao Report) at ¶ 13	Undisputed.
	Fact 20: Anthropic co-founder and Chief Science Officer, Dr. Jared Kaplan, attested: "To build an LLM like Claude, engineers first develop a 'neural network,' which is a computer program capable of studying enormous datasets of text and extracting statistical and linguistic information about those texts in an effort to gain a generalized understanding of human language." He also attested: "The neutral network is shown a vast dataset consisting of hundreds of millions, if not billions, of pieces of content. This set of materials is	Undisputed that Dr. Kaplan attested as quoted.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	known collectively as the 'training corpus' or simply the 'corpus.'" Kaplan PI Decl. ¶¶ 17, 18	
	Fact 21: Anthropic co-founder and Chief Science Officer, Dr. Jared Kaplan, attested: "The neural network then analyzes each token in the context of surrounding tokens to learn language patterns and the relationship between the words and phrases in the corpus. Through iterative training on large amounts of data, the neural network discovers patterns among words and phrases, extracting statistical information to develop a probabilistic but comprehensive map of how language works, without explicit human-defined rules. The model stores this statistical information in an organized set of numerical values called parameters, which include weights and biases (representing the connective strengths among neurons)." Kaplan PI Decl. ¶ 20	Undisputed that Dr. Kaplan attested as quoted.
	Fact 22: Publishers' expert, Dr. Shawn Shan, reported: "A large and diverse training corpus exposes the LLM to a variety of language patterns, allowing the LLM to learn from and model diverse sentence structures, topics, grammar, and other nuances. Training on diverse datasets provides models with more complete and comprehensive probabilities for different token patterns representing data and information in different contexts, which generally lead to more flexible, useful, and coherent outputs."	Undisputed.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	Ex. 7 (Shan Report) at ¶ 21; Ex. 131 (Shan Dep.) at 125:10–128:6 (Dr. Shan's testimony affirming the accuracy of the above statement from his report).	
	Fact 23: Anthropic's expert, Dr. Jeffrey Bilmes, reported: "Anthropic uses training data to train its models for the purpose of mapping out statistical relationships between text fragments so that the LLM models can generate new texts." Ex. 2 (Bilmes Report) at ¶ 36	Undisputed that Dr. Bilmes' report contains the quoted language. Anthropic relies on this fact for the truth of the matter asserted by the quoted language. Publishers dispute the truth of the matter asserted that Anthropic trains its models for the purpose of mapping out statistical relationships between text fragments so that the LLM models can generate new texts. The goal of an LLM is to accurately reproduce patterns and information from its training data while generating coherent textual responses to inputs that seem like the most reasonable response, as defined by statistics present in the training data. **Burrell Decl., Ex. 10** (Zhao Report) at ¶ 26. The goal of the LLM training process is to compress the training data in such a way as to enable the LLM to accurately reproduce the data when prompted. **Burrell Decl., Ex. 10** (Zhao Report) at ¶ 43; *see also* **Pls.' Opp'n/Reply Ex. 154** (Dep. Ex. 188, Anthropic_0000012356) at -360. In the LLM research community, it is still uncertain whether an LLM is even capable of generating "new" texts (as opposed to memorizing and regurgitating its training data).

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
		Zhao Decl. ¶¶ 44–47; **Burrell Decl., Ex. 10** (Zhao Report) ¶¶ 26, 46, 91.
	Fact 24: Anthropic's expert, Dr. Jeffrey Bilmes, reported: "[W]hile LLMs can perform complex tasks such as summarizing information and generating code, they do not function similarly to databases or search engines." And that "LLMs do not store or retrieve passages of text verbatim. Instead, they generate new and different outputs, probabilistically, based on general statistical patterns and relationships learned during training that are represented in the model's weights. Because their outputs are drawn from probability distributions, LLMs can produce different results when generating outputs in response to the same prompt, particularly when the final text is produced by randomly sampling from those distributions, which is by far the most common approach to LLM generation." Ex. 2 (Bilmes Report) at ¶ 98	Undisputed that Dr. Bilmes' report contains the quoted language. Anthropic relies on this fact for the truth of the matter asserted by the quoted language. Publishers dispute the truth of the matter asserted that LLMs do not function similarly to databases or search engines. LLMs store their training datasets in tokenized format, along with conditional probabilities that instruct the model to re-integrate the tokenized data in response to user queries, effectively functioning as a database that stores the training data text. **Burrell Decl., Ex. 10** (Zhao Report) at ¶¶ 21, 42, 73; *see also* **Pls.' Ex. 25** (Def.'s Am. Suppl. Resp. to Pls.' Interrog. No. 9, Nov. 12, 2025), at 8. Publishers dispute the truth of the matter asserted that LLMs do not store or retrieve passages of text verbatim. LLM training results in the LLM memorizing and storing highly-likely conditional probabilities from its training data that a particular sequence of tokens follows another particular sequence of tokens, resulting in the reproduction of specific phrases or texts from the LLM's training data verbatim. **Burrell Decl., Ex. 10** (Zhao Report), at ¶¶ 26–28, 43, 63; **Burrell Decl., Ex. 18** (Zhao Rebuttal) at ¶¶ 9–12, 22; **Pls.' Opp'n/Reply Ex. 154** (Dep. Ex. 188, Anthropic_0000012356) at -360; **Pls.' Opp'n/Reply Ex. 152** (Dep. Ex. 137, Anthropic_0000520123) at -123–25; **Pls.' Opp'n/Reply Ex. 136** (Drain

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
		Dep. Tr.) at 60:11–62:20 (testifying about Dep. Ex. 137). Publishers dispute the truth of the matter asserted that LLMs generate "new and different outputs." In the LLM research community, it is still uncertain whether an LLM is even capable of generating "new" content (as opposed to memorizing and regurgitating its training data). **Zhao Decl.** ¶¶ 44–47; **Burrell Decl., Ex. 10** (Zhao Report) ¶¶ 26, 46, 91.
	Fact 25: Publishers' expert, Dr. Ben Zhao, testified that "one of the purpose of lyrics is—and one of its primary intended use purposes is to enjoy words to a song." He answered, "I guess, in general, yes" to the question "When you listen to a piece of music, is the intended purpose for listening to that piece of music to enjoy it?" Ex. 139 (Zhao Dep.) at 352:23-353:16	Undisputed.
	Fact 26: Music compositions communicate the artistic expression of a musician or songwriter with an audience and to entertain. Ex. 137 (Stokes Dep.) at 230:11-18, 231:8-232:5; Ex. 26 (Stokes Reply) at ¶¶ 13-19	Undisputed to the extent the asserted fact accurately describes one purpose of musical compositions. Disputed to the extent the asserted fact reflects Mr. Stokes' disputed opinion, which is subject to a pending motion to exclude. *See* ECF No. 652.
	Fact 27: Publishers' expert, Dr. Ben Zhao, testified that Claude's "training corpus is decoupled from the model once the model is trained"; that Claude does not "stor[e] an exact copy of that text"; and that "the model does not scan for the text from its training corpus." Ex. 139 (Zhao Dep.) at 177:2-4,	Disputed. The asserted fact does not accurately reflect Dr. Zhao's testimony. In response to the question, "And so using the common computer science dictionary—or computer science definition of 'database,' LLMs do not store data or text in a database and retrieve that data, correct?", Dr. Zhao responded in full, "By that very strict

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	178:6-18, 179:9-10	definition of database, that is true, but I will just remark that computers in general, very few computers actually rely on [an] internal database. In fact, almost every piece of software that we use today does not include a formal database internally. That does not mean that it does not store content. That does not mean that it does not transform and store different versions of the content in different representative forms. You know, yeah, the idea that a strict definition of a database, yes, I agree that the LLMs do not store a strict definition of a database, but they are accessing different means and different representations." **Burrell Decl., Ex. 139** (Zhao Dep.) at 176:10–177:5. In response to the question "You say [in your rebuttal report]: 'The model has not memorized the specific text in the sense of literally storing an exact copy of that text and it does not strictly refer back to and fetch that sequence from its training corpus when it regurgitates. A pre-trained model does not scan for text from its training corpus and reproduce what it finds when asked.'. . . [d]o you stand by that statement?", Dr. Zhao responded, "Yes, the model does not scan for the text from its training corpus and produce it exactly." **Burrell Decl., Ex. 139** (Zhao Dep.) at 178:2-18. In response to the question, "When a user prompts a model like Claude, the model does not go to its training corpus, retrieve that text, and then reproduce or deliver that text to the—

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
		to the user, correct?", Dr. Zhao testified, "That is correct in the sense that I believe the training corpus is decoupled from the model once the model is trained," and clarified that, "Under the interpretation that the model parameters are, in fact, a lossy compressed version of the training corpus then, you know, it would have access, just in a slightly different format, a different representation format." **Burrell Decl., Ex. 139** (Zhao Dep.) at 179:2-24.
	Fact 28: Publishers' expert, Dr. Ben Zhao, reported "The model has not memorized the specific text in the sense of literally storing an exact copy of that text." Ex. 18 (Zhao Rebuttal) at ¶ 11	Disputed. The asserted fact does not accurately reflect Dr. Zhao's report. The full paragraph states: "As discussed in my initial report, conditional probability training results in LLMs memorizing unique phrases or sequences of phrases that do not commonly occur across multiple training samples, such as specific turns of phrases found in song lyrics. This results in regurgitation, wherein a model reproduces a specific text sequence in the same order that matches text in its training corpus. The model has not memorized the specific text in the sense of literally storing an exact copy of that text, and it does not strictly refer back to and fetch that sequence from its training corpus when it regurgitates. A pre-trained model does not scan for text from its training corpus and reproduce what it finds when asked. However, LLMs do memorize highly-likely conditional probabilities based on their training corpora (*e.g.*, the sequence of tokens 'Drove my Chevy to the levee, but the levee was dry' very- and highly-likely follows the sequence 'bye-bye, Miss

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
		American Pie'), which results in the same outcome: the reproduction of specific phrases or entire texts contained in its training corpus. This memorization of highly-likely conditional probabilities is fundamental to the transformer architecture of LLMs." **Burrell Decl., Ex. 18** (Zhao Rebuttal) at ¶ 11.
	Fact 29: Anthropic's expert, Dr. Jeffrey Bilmes, reported: "Anthropic uses training data to train its models for the purpose of mapping out statistical relationships between text fragments so that the LLM models can generate new texts." Ex. 2 (Bilmes Report) at ¶ 36	Undisputed that Dr. Bilmes' report contains the quoted language. Anthropic relies on this fact for the truth of the matter asserted by the quoted language. Publishers dispute the truth of the matter asserted that Anthropic trains its models for the purpose of mapping out statistical relationships between text fragments so that the LLM models can generate new texts. The goal of an LLM is to accurately reproduce patterns and information from its training data while generating coherent textual responses to inputs that seem like the most reasonable response, as defined by statistics present in the training data. **Burrell Decl., Ex. 10** (Zhao Report) at ¶ 26. The goal of the LLM training process is to compress the training data in such a way as to enable the LLM to accurately reproduce the data when prompted. **Burrell Decl., Ex. 10** (Zhao Report) at ¶ 43; *see also* **Pls.' Opp'n/Reply Ex. 154** (Dep. Ex. 188, Anthropic_0000012356) at -360.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
		In the LLM research community, it is still uncertain whether an LLM is even capable of generating "new" texts (as opposed to memorizing and regurgitating its training data). **Zhao Decl. ¶¶** 44–47; **Burrell Decl., Ex. 10** (Zhao Report) ¶¶ 26, 46, 91.
	Fact 30: Anthropic's expert, Dr. Jeffrey Bilmes, reported: "[W]eights are one type of model parameter for an NN [neural network]. Weights are the adjustable values in the model that model training seeks to optimize. Weight tuning is a subprocess of model training where the weights of an NN are adjusted iteratively using the training data so that the model's predictions better align with the expected outputs." Ex. 2 (Bilmes Report) at ¶ 52	Undisputed that Dr. Bilmes' report contains the quoted language.
	Fact 31: Tokenizing text breaks it into small, standard units, identified numerically, that a large language model can process. Tokens are then transformed into richer numerical representations in "embeddings," which encode a token's meaning and semantic relationship to other tokens. Token embeddings are learned parameters that evolve during model training, and as the model processes billions of data points, mathematically move closer to other tokens that appear in similar contexts. Embeddings capture the relationship between words, making mathematical reasoning about those words' meanings possible. Embeddings are then positionally encoded so that a model can recognize both what each word means and where it occurs in context, rather than	Disputed. A trained LLM stores the conditional probabilities (the likelihood of a particular sequence of tokens following other sequences of tokens) found in its training data. This results in memorization of unique phrases found in the model's training data, which in turn results in regurgitation of specific text sequences in the same order found in the training data, thus "stor[ing] text verbatim." **Burrell Decl., Ex. 10** (Zhao Report) ¶¶ 13, 28; **Burrell Decl., Ex. 18** (Zhao Rebuttal) ¶ 11; **Burrell Decl., Ex. 139** (Zhao Dep. Tr.) at 176:10–180:19.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	simply storing text verbatim. Tokens, embeddings, and positional encoding allow models to understand the contextual relationship of words represented numerically. Ex. 2 (Bilmes Report) at ¶¶ 141-152	
	Fact 32: Publishers' expert, Dr. Ben Zhao, reported: "Through pre-training, the model will learn which of these is the best choice from context. Pre-training also involves 'weighting' the dataset. In this process, the trainer directs the model to use some parts of the dataset more often than others due to differences in perceived quality of the material, with portions engaged multiple times and others not at all. Ultimately, the model will learn the most common patterns in the grammar, syntax, and semantics of a given language. This process requires no supervision (i.e., labeling or metadata describing the contents of the text). Rather, the model recognizes patterns in the training data to infer these rules automatically." Ex. 10 (Zhao Report) at ¶ 20	Undisputed.
	Fact 33: Publishers' expert, Dr. Ben Zhao, reported, "Claude, like all LLMs, does not rely on an inherent understanding of truth or facts in the same way humans do." And that Anthropic has trained Claude to "generate[] output based on learning statistical associations between tokenized text." Ex. 10 (Zhao Report) at ¶ 141	Undisputed.

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	Fact 34: Publishers' expert, Dr. Shawn Shan, testified that "[i]t is possible" that "there are going to be instances in which when it's ask[ed] for something that's not very simple, Claude will output different outputs in response to identical prompts, even at the same time." Ex. 131 (Shan Dep.) at 247:7-13	Undisputed.
	Fact 35: LLMs process a user's prompt and draw on learned language patterns to stochastically output the next predicted word. Model builders control the level of stochasticity (or randomness) in the output, including through a parameter called the "temperature parameter." As the temperature increases, so does the randomness of the output. This is unlike a database or retrieval system, which deterministically fetch and return the same result for the same query. This makes it possible for LLM's to generate outputs that are linguistically and textually different while containing similar factual content. Ex. 2 (Bilmes Report) at ¶¶ 271, 280-285 & Figure 33	Disputed. A trained LLM stores the conditional probabilities (the likelihood of a particular sequence of tokens following other sequences of tokens) found in its training data. This results in memorization of unique phrases found in the model's training data, which in turn results in regurgitation of specific text sequences in the same order found in the training data; in that way, the model "fetch[es] and return[s]" text from its training data and reproduces such text. **Burrell Decl., Ex. 10** (Zhao Report) ¶¶ 13, 28; **Burrell Decl., Ex. 18** (Zhao Rebuttal) ¶ 11; **Burrell Decl., Ex. 139** (Zhao Dep. Tr.) at 176:10–180:19.
	Fact 36: Publishers' expert, Dr. Ben Zhao, reported that "The model has not memorized the specific text in the sense of literally storing an exact copy of that text, and it does not strictly refer back to and fetch that sequence from its training corpus when it regurgitates. A pretrained model does not scan for text from its training corpus and reproduce what it finds when asked." When asked in his deposition whether he "stand[s] by	Undisputed that Dr. Zhao testified as quoted during his deposition. Otherwise, disputed. The asserted fact does not accurately reflect Dr. Zhao's report. The full paragraph states: "As discussed in my initial report, conditional probability training results in LLMs memorizing unique phrases or sequences of phrases that do not commonly occur across multiple training samples, such as specific turns of phrases found in song lyrics.

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	that statement," Zhao answered, "Yes, the model does not scan for the text from its training corpus and produce it exactly." Ex. 18 (Zhao Rebuttal) at ¶ 11; Ex. 139 (Zhao Dep.) at 178:4-18	This results in regurgitation, wherein a model reproduces a specific text sequence in the same order that matches text in its training corpus. The model has not memorized the specific text in the sense of literally storing an exact copy of that text, and it does not strictly refer back to and fetch that sequence from its training corpus when it regurgitates. A pre-trained model does not scan for text from its training corpus and reproduce what it finds when asked. However, LLMs do memorize highly-likely conditional probabilities based on their training corpora (*e.g.*, the sequence of tokens 'Drove my Chevy to the levee, but the levee was dry' very- and highly-likely follows the sequence 'bye-bye, Miss American Pie'), which results in the same outcome: the reproduction of specific phrases or entire texts contained in its training corpus. This memorization of highly-likely conditional probabilities is fundamental to the transformer architecture of LLMs." **Burrell Decl., Ex. 18** (Zhao Rebuttal) at ¶ 11.
	Fact 37: Anthropic displays a webpage titled, "Build With Claude," which markets "Use Cases for Claude" to enterprise clients. Those use cases include: "Coding" where its "latest model, Claude Opus 4.1, can complete complex engineering tasks to solve problems that would typically take a full day of engineering work"; "Agents," offering "superior instruction following, tool selection, error correction, and advanced reasoning for customer-facing agents and complex AI workflows"; "Productivity" where "Claude can	Undisputed that the cited document, dated September 8, 2025, contains the quoted language.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	extract relevant information from business emails and documents, categorize and summarize survey responses, and wrangle reams of text with high speed and accuracy"; and "Customer Support." Ex. 63 (Anthropic_0000437208)	
	Fact 38: Anthropic's 30(b)(6) witness, Dr. Deep Ganguli, testified: "[T]he main use case of CLAUDE.AI and CLAUDE Code and our CLAUDE API is for writing code. It's for doing software engineering tasks. That is the most common use case across—across the product surfaces I mentioned." Ex. 119 (Ganguli 30(b)(6) Dep.) at 67:7-15	Undisputed that Dr. Ganguli testified as quoted.
	Fact 39: Publishers' expert, Dr. Shawn Shan, testified: "LLMs, one of the use case is, as we describe, a user go to Claude.ai or .com to use the service but if I understand correctly, that is a fraction of how LLMs are used today" and "Claude is also used as a reasoning systems or an agent, where, you know, user ask original task, but Claude will generate command and prompt itself to do a step-by-step educate—execution of the task." Ex. 131 (Shan Dep.) at 95:22-96:17	Undisputed that Dr. Shan testified as quoted on March 3, 2026.
	Fact 40: Anthropic's 30(b)(6) witness, Jane Leibrock, testified: "We promote CLAUDE's safety as a model, its accuracy, its high intelligence, its skill at working with a large amount of knowledge by which we mean data that's been provided by the user to the model, its ability to show its chain of thought or reasoning to do deep	Undisputed that Ms. Leibrock testified as quoted.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	research, and to work with software code or to debug code." Ex. 122 (Leibrock Dep.) at 26:14-27:4	
	Fact 41: Anthropic co-founder and Chief Executive Officer, Dario Amodei, testified: "Anthropic, it's very focused on code. It's very focused on work and productivity-based activities. So Anthropic has deliberately made—made choices not to—not to produce output in many creative modalities." *See* Ex. 112 (Dario Amodei Dep.) at 102:19-24	Undisputed that Dr. Amodei testified as quoted. Anthropic relies on this fact for the truth of the matter asserted by the quoted testimony. Publishers dispute the truth of the matter asserted that Anthropic has deliberately made choices not to produce output in many creative modalities. When Anthropic released Claude to the public in March 2023, Anthropic specifically identified "use cases including summarization, search, ***creative and collaborative writing***, Q&A, coding, and more." **Pls.' Ex. 73** (Dep. Ex. 187, Anthropic_0000349548) (emphasis added); *see* **Pls.' Ex. 4** (Mann 30(b)(6) Dep. Tr.) at 21:22–22:6, 22:22–23:3, 23:8-10, 26:1-4. Publishers further dispute the truth of the matter asserted that Anthropic has deliberately made choices not to produce output in many creative modalities because Claude is frequently used for a variety of such use cases, including "[w]rit[ing] words to fit musical compositions, including lyrics" and "[w]rit[ing] or improv[ing] rap lyrics." **Pls.' Ex. 105** (D. Ganguli 30(b)(6) Dep. Ex. 7, Anthropic_0000441309); **Pls.' Ex. 30** (D. Ganguli 30(b)(6) Dep. Ex. 6, Anthropic_0000429489) at Rows 2620, 3044, 4271, 4474, 4758; *see also* Smith Decl. ¶ 46.
	Fact 42: Anthropic co-founder and President, Ms. Daniela Amodei,	Undisputed that Ms. Amodei testified as quoted on November 12, 2025.

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	testified that Anthropic "serve[s] about 300,000 businesses." Ex. 111 (Daniela Amodei Dep.) at 68:7-9	Anthropic relies on this fact for the truth of the matter asserted by the quoted testimony. Publishers dispute the truth of the matter asserted that Anthropic serves 300,000 businesses. The number of Anthropic's business customers has grown significantly since the time of this testimony. **Pls.' Ex. 96** (ANTHROPIC, *Anthropic raises $30 billion in Series G funding at $380 billion post-money valuation* (Feb. 12, 2026)) ("Business subscriptions to Claude Code have quadrupled since the start of 2026").
	Fact 43: Anthropic co-founder, Mr. Tom Brown, testified "we want Claude to be the best model for people who want to do work, particularly coding work." He also testified "We have—we have really focused on work rather than the consumer application" of Claude. He testified "we have always wanted to prioritize data that makes the model good at work." Ex. 106 (Brown Dep.) at 48:17-21, 64:4-8, 64:12-16, 64:23-65:4, 86:3-7, 94:11-16	Undisputed that Mr. Brown testified as quoted. Anthropic relies on this fact for the truth of the matter asserted by the quoted testimony. Publishers dispute the truth of the matter asserted that Anthropic has focused on work rather than the consumer application, because Anthropic specifically identified "use cases including summarization, search, ***creative and collaborative writing***, Q&A, coding, and more" when it first launched Claude in March 2023. **Pls.' Ex. 73** (Dep. Ex. 187, Anthropic_0000349548) (emphasis added); *see* **Pls.' Ex. 4** (Mann 30(b)(6) Dep. Tr.) at 21:22–22:6, 22:22–23:3, 23:8-10, 26:1-4. Publishers further dispute the truth of the matter asserted that Anthropic has "focused on work rather than the consumer application," because Anthropic targets "ordinary retail users" as part of its marketing.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
		Burrell Decl., Ex. 122 (Leibrock Dep) at 21:10-17.
	Fact 44: Anthropic co-founder Tom Brown used Claude to explore how to write his wife a song for Valentine's Day given his wife's taste for Adele's music. In the course of the conversation, Claude explained: "Since you mentioned Adele's influence, we could focus on emotional storytelling and specific details that make it personal to your relationship. Adele often structures her songs with … [i]ntimate verses that tell a story[,] [a] build-up to an emotional chorus[,] [b]ridge sections that add depth[,] [and] [p]owerful vocal movements." Ex. 95 (Anthropic_0000570705)	Undisputed.
	Fact 45: Anthropic co-founder Tom Brown asked Claude, "What are your favorite pearl jam tracks?" In response, Claude listed five songs, including "Black." Claude explained that "Black" "captures that devastating moment when you realize someone you love has moved on completely while you're still consumed by the relationship," and "[t]he raw emotion in Vedder's voice during the outro ('I know someday you'll have a beautiful life...') creates one of rock's most heartbreaking moments." Ex. 92 (Anthropic_0000570614)	Undisputed.
	Fact 46: Claude has recommended and discussed musical compositions with users based on their music taste. Anthropic's President, Ms. Daniela Amodei, asked Claude for suggestions of classical music that is "meaningful but positive / uplifting," Claude responded with a	Undisputed.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	list of ten musical works. It then explained: "Beethoven's 'Ode to Joy' is a powerful anthem of brotherhood and joy. Handel's 'Hallelujah Chorus' is jubilant and celebratory. Mozart's 'Eine kleine Nachtmusik' is light and cheerful, while Vivaldi's 'Spring' from The Four Seasons evokes the rebirth and energy of springtime." Ex. 96 (Anthropic_0000570712) at -713	
	Fact 47: Anthropic co-founder, Mr. Benjamin Mann, asked Claude "What's the story behind the song dancing in the moonlight" and Claude explained that "The song was originally written and recorded by French-American songwriter and musician Sherman Kelly in 1969 when he was part of a band called Boffalongo." Claude continued that "The song emerged as a direct contrast to the violence [Kelly] had experienced" "while vacationing on the Caribbean island of St. Croix" to "create[] a vision of a peaceful, harmonious world where people came together in joy rather than violence." "The uplifting, almost utopian lyrics of the song ('Everybody's dancing in the moonlight…') represents his yearning for a more peaceful existence after facing such brutality." Ex. 90 (Anthropic_0000570582)	Undisputed.
	Fact 48: Anthropic technical staff, Mr. Ben Kuhn, asked Claude for "some good three-part choral music for two friends" and Claude "recommend[ed] some three-part choral music that would work well for sightreading with friends"	Undisputed.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	including "Pelstrina's 'Sicut Cervus'" "Orlando di Lasso's 'Matona, mia cara'" "Schubert's 'Sanctus'" "Avro Pärt's simpler works like 'Morning Star'" "Eric Whitacre's 'Seal Lullaby'" and more, categorizing them by "different difficulty levels and styles." Ex. 89 (Anthropic_0000570574)	
	Fact 49: Anthropic co-founder, Mr. Tom Brown, asked Claude about Jimi Hendrix and the "wah wah" pedal, and Claude responded that "Hendrix helped popularize the wah-wah pedal in the late 1960s, using it to great effect on songs like 'Voodoo Child (Slight Return)' and 'All Along the Watchtower.'" Claude continued that "The pedal essentially became an extension of the instrument, allowing players to 'voice' their guitar in new ways." Ex. 94 (Anthropic_0000570704)	Undisputed.
	Fact 50: Anthropic co-founder, Mr. Tom Brown, asked Claude about Joni Mitchell's "Both Sides Now" and why he "like[s] it despite it being something that I wouldn't have liked in my twenties." Claude responded that "The emotional complexity and introspective lyrics of 'Both Sides Now' might resonate more deeply now that you've had more life experience. The song is fundamentally about seeing things from multiple perspectives and understanding life's contradictions—themes that often become more meaningful as we gain life experience." When Mr. Brown asked why he "might have dismissed [other music] which are actually good[] Especially things	Undisputed.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	from the 80s?" Claude responded that "The 1980s are particularly interesting because they often get dismissed for their production style (lots of synthesizers, gated reverb on drums, sometimes 'cheesy' by modern standards) or theatrical presentation" before listing artists like "Kate Bush" "Peter Gabriel" "Tears for Fears" and "The Eurythmics" that bore thematic similarities to Joni Mitchell. The conversation then discusses more modern artists whose "content and wisdom" align with Mitchell's and Claude discusses how Kendrick Lamar's lyrics in " 'Mother I sober' where he says 'I'm sensitive, I feel everything, I feel everybody / One man standing on two words, healing and hurt' – which suggests he's found some path toward healing while acknowledging the hurt is still there." Other Anthropic employees asked Claude to: • help put together a book of song lyrics, • explain the meaning of Stephen Sondheim's "Pacific Overtures," • explain the meaning behind the song "Every Breath You Take by the Police," to which Claude responded "The lyrics … reveal the narrator's controlling and stalking behavior toward someone who has left them…. The song's dark themes are somewhat masked by its melodic, gentle musical arrangement." • explain the story behind	

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	Bob Dylan's "Mr. Tambourine Man" and Cat Steven's "Wild World." • to ask for "some song to sing" to a "newborn son" because "m[y] dad sang James Taylor's 'Carolina on my mind'" and Claude gave recommendations on "wonderful options for singing to you son." Ex. 93 (Anthropic_0000570681); Ex. 147 (Anthropic_0000570572); Ex. 148 (Anthropic_0000570592); Ex. 149 (Anthropic_0000570627); Ex. 150 (Anthropic_0000570635); Ex. 151 (Anthropic_0000570636)	

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	Publishers' Additional Fact #29: Memorization of portions of the training data is an inherent trait of LLM training. **Burrell Decl. Ex. 10** (Zhao Report) at ¶¶ 26, 43, 63; **Pls.' Opp'n/Reply Ex. 136** (Drain Dep. Tr.) at 69:15–71:7; **Pls.' Opp'n/Reply Ex. 135** (Sangani Dep. Tr.) at 172:5-12 ("I would not frame regurgitation as a risk. I think the risk is memorization. And if a large contiguous block of tokens has been memorized, it's—it's possible that for some combination of prompts, those tokens will be reproduced. Maybe they would not be regurgitated; maybe they would. It's hard to say, but it's not unreasonable that that could occur"); *see also* Answer to FAC ¶¶ 108–09; **Pls.' Opp'n/Reply Ex.**	Disputed that memorization is inherent or desirable. "Generalization is fundamentally distinct from, and opposed to, memorization. Model builders, like Anthropic, aim for their models to be able to generalize, not memorize." Ex. 2 (Bilmes Report) ¶¶ 87, 117-120. Anthropic in fact takes substantial steps to mitigate memorization, including through deduplicating its training data because memorization hurts model performance. Pls.' Opp'n/Reply Ex. 136 (Drain Dep.) at 70:13-18; Ex. 116 (Drain Dep.) at 66:9-23, 83:14-18; 154:5-15, 178:17-179:4, 229:8-18; Ex. 2 (Bilmes Report) ¶¶ 87, 117-120; Ex. 126 (Mann 30(b)(6) Dep.) at 178:9-14; Ex.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	129 (Brown Dep. Tr.) at 196:1-4 ("I think that if you trained a ton on—like, in pretraining some verbatim chunk of text many, many times, that that would make it more likely that the model would regurgitate it.").	112 (Dario Amodei Dep.) at 94:22-25; Ex. 106 (Brown Dep.) at 131:8-17. Further disputed as the cited testimony from Mr. Pranay Sangani and Mr. Tom Brown concerned memorization of portions of the training data as a *hypothetical* trait, not an "inherent" one. In the cited testimony, Mr. Pranay Sangani was responding to the question, "[I]f a model is engaged in memorization, does that increase the risk of potential regurgitation?" Pls.' Opp'n/Reply Ex. 135 (Sangani Dep.) at 172:1-12. Mr. Tom Brown was responding to the question, "And is it your testimony that you don't -- that you can't say one way or the other if specific content is repeated many times in the training data, whether that's more likely to influence the model's output? Mr. Brown's full response was, "I think I don't -- I think I don't know. I think that if you trained a ton on -- like, in pretraining some verbatim chunk of text many, many times, that that would make it more likely that the model would regurgitate it. But again, I think that's – that's the reason why we avoid doing stuff like that, but again, I'm not the expert here. Ant. Reply Ex. 165 (Brown Dep.) at 195:19-197:7.
	Publishers' Additional Fact #30: [INTENTIONALLY OMITTED]	No reply required.
	Publishers' Additional Fact #31: Other top use cases of Claude include "helping people to actually do work, whether that's writing, whether that's doing research, or whether that's an educator doing	Undisputed that Dr. Deep Ganguli testified to the quoted language and that Dr. Deep Ganguli further testified that "at the very high level, the main use case of CLAUDE.AI and CLAUDE Code and our CLAUDE

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	kind of curriculum design," as Claude is "a very general-purpose technology." **Burrell Decl., Ex. 119** (Ganguli 30(b)(6) Dep.) at 67:7-24.	API is for writing code. It's for doing software engineering tasks. That is the most common use case across – across the product surfaces I mentioned." And that the other use cases cited are "the next kind of top uses". Ex. 119 (Ganguli 30(b)(6) Dep.) at 67:7-24.
	Publishers' Additional Fact #32: Anthropic trains Claude to generate preferred outputs and avoid outputs Anthropic deems undesirable. Answer to FAC ¶ 58(d); Kaplan Decl. ¶¶ 38–41; **Pls.' Ex. 4,** Mann 30(b)(6) Dep. Tr. 149:13–150:8; **Pls.' Ex. 120**, Yuntao Bai et al., *Training a Helpful and Harmless Assistant with Reinforcement Learning from Human Feedback*,ARXIV:2204.05862v1, 4 (Apr. 12, 2022), https://arxiv.org/pdf/2204.05862 (describing the process of using human feedback data to train an AI model to understand which responses are harmful and which responses are helpful and honest); **Kaplan PI Decl.** ¶ 38 ("Constitutional AI is at the heart of all of Anthropic's models, and is designed to ensure AI safety by fine-tuning the models on a set of principles that Anthropic believes will cultivate steerable, helpful model behavior that aligned with human preferences."); **Pls.' Ex. 25** (Def.'s Am. Suppl. Resp. to Pls.' Interrog. No. 12, Nov. 12, 2025) at 12.	Undisputed that Anthropic trains Claude to generate helpful, honest, and harmless outputs including through finetuning. It is further undisputed that Anthropic utilizes reinforcement learning with human feedback, among other finetuning techniques, to align Claude to human preferences of helpfulness, honesty, and harmlessness. And further undisputed that Anthropic trains Claude to avoid reproducing copyrighted materials in outputs. Kaplan PI Decl. ¶ 38; Otherwise disputed to the extent that Publishers imply Anthropic can dictate or predict specific outputs. Anthropic does not control Claude's exact outputs, which are stochastic (i.e., random and variable). Claude's outputs are not deterministically generated. LLMs, like Claude, generate outputs randomly and variably by sampling from probability distributions in model weights, meaning there is always some randomness and variability in Claude's outputs. Ex. 2 (Bilmes Report) ¶¶ 86, 114, 198-201, 271, 281-282; Ex. 11 (Bilmes Rebuttal) ¶¶ 26, 45, 99, 102.
	Publishers' Additional Fact #33: In response to Anthropic founder Tom Brown's prompts seeking	Undisputed that the cited document contains the quoted text.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	assistance in writing a song for his wife, Claude offered to start by brainstorming lyrics: "[W]e could start by brainstorming lyrics that capture your feelings and relationship, taking inspiration from the emotional depth of Adele's songwriting." **Burrell Decl., Ex. 95** (Anthropic_0000570705), -705.	Otherwise disputed that Claude outputted Adele lyrics in response to Mr. Tom Brown's request. Claude stated, "*Since you mentioned Adele's influence*, we could focus on emotional storytelling and specific details that make it personal to your relationship. Adele often structures her songs with [i]ntimate verses that tell a story[,] [a] build-up to an emotional chorus, [b]ridge sections that add depth, [and] [p]owerful vocal moments." Ex. 95 (Anthropic_0000570705), -705.
	Publishers' Additional Fact #34: In response to Anthropic founder Tom Brown's question, "Does Adele-style music tend to rhyme?", Claude generated output containing portions of the lyrics to two Adele songs, "Someone Like You" and "Hello." **Burrell Decl., Ex. 95** (Anthropic_0000570705), -706.	Undisputed that Tom Brown prompted Claude with the question "Does Adele-style music tend to rhyme?" Disputed as to the truth of the matter asserted. Mr. Tom Brown testified that he was planning to write an original song for his wife for Valentine's Day. He used Claude to coach him through the process. Mr. Tom Brown identified Adele as one of his wife's favorite artists and asked whether her music tends to rhyme. Claude responded by explaining that Adele "uses a natural, conversational style that doesn't feel forced." There is no evidence that the purported lyrics to "Someone Like You" or "Hello" in Claude's output are accurate and neither are works in suit. Ex. 95 (Anthropic_0000570705), -706; Ex. 106 (Brown Dep.) at 129:22-131:7; Dkt. 337-1 (Ex. A to FAC).
	Publishers' Additional Fact #35: Claude is capable of generating output regarding song meaning and musical artists because "that's something that people talk about a	Undisputed that "Claude is capable of generating output regarding song meaning and musical artists" and that these are topics "people talk about a lot on the Internet" and that there is a

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	lot on the Internet," and there is "a wide variety of publicly available commentary" on songs on the internet that Anthropic copies and includes in Claude's training data. **Pls.' Opp'n/Reply Ex. 129** (Brown Dep. Tr.) at 149:8–150:18, 181:2–182:7, 183:12-18, 193:1-14.	"wide variety of publicly available commentary on." It is further undisputed that Claude is trained on this "publicly available" text from "from the Internet." Disputed that the capabilities cited stem from these sources alone as LLM's generate outputs, probabilistically, based on general statistical patterns and relationships learned during training that are represented in the model's weights and do not fetch and retrieve data from specific sources in their training data. Ex. 2 (Bilmes Report) ¶¶ 36, 98, 271, 280-285 & Figure 33.
	Publishers' Additional Fact #36: Claude's output regarding songs can include discussion of auditory and acoustic sensory experience (such as a singer's "gravelly" voice or "crystalline vocal quality," or a melody including a "jangly ringing sound"), which is not based on Anthropic's training Claude on song lyrics or audio recordings, but rather on Anthropic's copying written commentary about songs. **Pls.' Opp'n/Reply Ex. 131** (Ganguli 30(b)(1) Dep. Tr.) at 102:15-18 ("Claude can't listen, especially not—not a 2023 version of this model"); **Burrell Decl., Ex. 93** (Anthropic_0000570681) at -681–82; **Burrell Decl., Ex. 150** (Anthropic_0000570635); **Pls.' Opp'n/Reply Ex. 129** (Brown Dep. Tr.) at 193:8–194:8; *see also* **Burrell Decl., Ex. 149** (Anthropic_0000570627); **Burrell Decl., Ex. 151** (Anthropic_0000570636) at -636–37.	Undisputed that Ex. 93 (Anthropic_0000570681) describes a singer's "gravelly" voice and another singer's "crystalline vocal quality. Undisputed that Claude is not trained on audio recordings. Undisputed that Ex. 150 (Anthropic_0000570635) states that "[t]he Byrds' arrangement transformed Dylan's acoustic folk song into something entirely new - they added electric guitars, harmonies, and a jangly, ringing sound that helped define what would become known as 'folk rock.'" Ex. 93 (Anthropic_0000570681) at -681–82; Ex. 150 (Anthropic_0000570635). Otherwise disputed that Claude's output is "based … on Anthropic's ***copying*** written commentary about songs." When asked how Claude is "capable of describing the 'flute melody as iconic and the beat as smooth and menacing,'" Mr. Tom

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		Brown testified: "My guess is that there's like a wide variety of publicly available commentary on this track and that Claude picked it up via learning about that commentary … My guess is that no person [has] used this exact sentence and it's more that Claude is, like, picking up kind of a vibe from the commentary and then synthesizing it into this summary." Pls.' Opp'n/Reply Ex. 129 (Brown Dep.) at 193:8–194:8.
	Publishers' Additional Fact #37: Anthropic does not need to train Claude on song lyrics in order for Claude to be able to write computer code or perform other technical tasks not related to song lyrics. **Pls.' Opp'n/Reply Ex. 131** (Ganguli 30(b)(1) Dep. Tr. 206:19–207:14); *see also* **Pls.' Opp'n/Reply Ex. 129** (Brown Dep. Tr.) at 85:6-15, 86:14-16, 88:1-3; **Pls.' Opp'n/Reply Ex. 138** (Altman Dep. Tr.) at 132:25–133:23, 141:25–142:19, 153:9-17, 165:1-13, 167:13–168:6, 173:7-20, 257:17–258:1 (expressing no opinion regarding the effect of removing song lyrics from Claude's training data on various purported benefits and capabilities of AI).	Disputed. Developing an LLM, like Claude, requires ingesting an enormous and diverse training corpus because the model learns language patterns by extracting statistical correlations across vast amounts of preexisting human writing. To assemble corpora of that scale, developers must rely on large-scale web crawling efforts that necessarily sweep in information that is publicly accessible on the internet. Each individual piece of training data has impact on the model's weights so each piece of training data has an impact on the model. Ex. 10 (Zhao Report) ¶ 13; Ex. 7 (Shan Report) ¶¶ 19, 21-23; Ex. 131 (Shan Dep.) at 125:10-128:6, 226:9 229:18; Ex. 2 (Bilmes Report) ¶¶ 100, 102; Ex. 6 (Nagaraj Report) ¶¶ 36-37; Dkt. 212-5, ("Kaplan PI Decl.") ¶¶ 29-30; Ex. 56 (Anthropic_0000231226).

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	Publishers' Additional Fact #38: Anthropic founder and Chief Compute Officer Tom Brown entered the prompt into Claude in February 2023 asking "@Claude what are the lyrics to [Work in Suit] desolation row by Dylan?" However, Mr. Brown "[didn't] really remember clearly" the output that Claude generated in response to that prompt. **Pls.' Opp'n/Reply Ex. 129** (Brown Dep. Tr.) at 210:17-23, 211:11–212:19, 213:3-10; Answer to FAC ¶¶ 11, 113; **Pls.' Ex. 58** (Dep. Ex. 274, Anthropic_0000016458).	Undisputed that the cited document contains the quoted text. Otherwise disputed as to the truth of the matter asserted. Mr. Tom Brown testified that "this is me playing around with my family like seeing what stuff the model could or could not do" and that "I think it refused the lyrics to the 'Desolation Row' one because I remember that was a question my dad asked. And then I think he was not impressed with – with CLAUDE at the time." Ex. 106 (Brown Dep.) at 209:11-210:11.
	Publishers' Additional Fact #39: Anthropic founder and President Daniela Amodei asked Claude to write a song to sing to her son, and Claude responded by generating purportedly new AI song lyrics. **Pls.' Ex. 110** (Dep. Ex. 174, Anthropic_0000570711); *see* **Pls.' Ex. 10,** Daniela Amodei Dep. Tr. 186:11-16.	Undisputed that Ms. Daniela Amodei asked Claude, "Can you write lyrics about a pig to the childrens [sic] song of baa baa black sheep for me to sing to my toddler?" Further undisputed that Claude responded "Oink oink pink pig, Playing in the hay, Rolling in the muddy path All throughout the day. Yes sir, yes sir, Having so much fun, One mad bath in morning time, And one when day is done." Pls.' Ex. 110 (Dep. Ex. 174, Anthropic_0000570711); *see also* Pls.' Ex. 10 (Daniela Amodei Dep.) at 186:11-16.
	Publishers' Additional Fact #40: Anthropic co-founder Benjamin Mann asked Claude to identify a Stevie Wonder song by prompting Claude with a lyric from the song. **Pls.' Ex. 85** (Dep. Ex. 215, Anthropic_0000524147); **Pls.' Ex. 5,** Mann 30(b)(1) Dep. Tr. 217:7-20.	Undisputed that Mr. Benjamin Mann asked Claude, "What song by Stevie Wonder includes the lyrics? Everyday's a New Day," and Claude responded, "I aim to be helpful while respecting copyright. The song you're asking about appears to be \'Each Day's A New Day\' (note the slight difference in the title compared to what was quoted). It was released by Stevie Wonder on his 1991 album \'Jungle Fever\' which was the soundtrack to the Spike Lee film of the same name. I apologize that I

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		cannot reproduce or confirm specific lyrics to avoid any copyright issues, but I can tell you it's an uplifting song about optimism and new beginnings.\n\nIf you'd like to find the official lyrics, I'd recommend checking authorized lyrics websites or Stevie Wonder's official materials. I'm happy to discuss other aspects of the song or album that don't involve reproducing copyrighted content." Pls.' Ex. 85 (Dep. Ex. 215, Anthropic_0000524147) at -147, -148
	Publishers' Additional Fact #41: Anthropic employees have prompted Claude with requests for lyrics to Publishers' Works and other songs. **Pls.' Ex. 54** (Anthropic_0000525947) at -948–49; **Pls.' Ex. 55** (Anthropic_0000459162) at -168; **Pls.' Ex. 56** (Anthropic_0000554366), at -387–88; **Pls.' Ex. 58** (Dep. Ex. 274, Anthropic_0000016458); **Pls.' Ex. 59** (Anthropic_0000016025) at -035; **Pls.' Ex. 60** (Anthropic_0000554319) at -322; **Pls.' Ex. 62** (Anthropic_0000527903) at -915; *see also* **Pls.' Ex. 80** (Anthropic_0000582187); **Pls.' Ex. 81** (Anthropic_0000430272); **Pls.' Ex. 82** (Anthropic_0000430273); **Pls.' Ex. 83** (Anthropic_0000430258); **Pls.' Ex. 84** (Anthropic_0000430285) at -286; Answer to FAC ¶¶ 11, 113.	Undisputed that Anthropic employees have prompted Claude with requests for lyrics. Insofar as Publishers Ad.F #41 relates to outputs, disputed because the cited exhibits either do not include the actual prompts submitted to Claude or do not demonstrate that Claude provides verbatim lyrics in response to requests for lyrics. The cited exhibits further demonstrate that Claude refuses to provide copyrighted output in response to prompts. These records were generally created before Anthropic augmented its guardrails intended to prevent the generation of outputs that may infringe copyrights and specifically the works in suit and Anthropic stipulated with Publishers to maintain their efficacy. Pls.' Ex. 54 (Anthropic_0000525947) at -948–49; Pls.' Ex. 55 (Anthropic_0000459162) at -168; Pls.' Ex. 56 (Anthropic_0000554366), at -387–88; Pls.' Ex. 58 (Brown Dep. Ex. 274, Anthropic_0000016458); Pls.' Ex. 59 (Anthropic_0000016025) at -035; Pls.' Ex. 60 (Anthropic_0000554319) at -322;

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		Pls.' Ex. 62 (Anthropic_0000527903) at -915; *see also* Pls.' Ex. 80; (Anthropic_0000582187); Pls.' Ex. 81 (Anthropic_0000430272); Pls.' Ex. 82 (Anthropic_0000430273); Pls.' Ex. 83 (Anthropic_0000430258); Pls.' Ex. 84 (Anthropic_0000430285) at -286; Pls.' Ex. 25 (Def.'s Am. Suppl. Resp. to Pls.' Interrog. No. 12, Nov. 12, 2025) at 11-14; Dkt. 291.
	Publishers' Additional Fact #42: Anthropic employees have prompted Claude with requests to identify full song lyrics based on submitted lyrics. **Pls.' Ex. 64** (Anthropic_0000459249) at -271; **Pls.' Ex. 65** (Anthropic_0000557304) at -328; **Pls.' Ex. 66** (Anthropic_0000459543) at -742; **Pls.' Ex. 67** (Anthropic_0000557359), -612–14; **Pls.' Ex. 79** (Anthropic_0000570595) at -596; **Pls.' Ex. 85** (Dep. Ex. 215, Anthropic_0000524147) at -147; **Pls.' Ex. 86** (Anthropic_0000431481); **Pls.' Ex. 87** (Anthropic_0000430104); **Pls.' Ex. 88** (Anthropic_0000430107); **Pls.' Ex. 89** (Anthropic_0000434264).	Disputed that Anthropic employees have prompted Claude with requests to identify "full song lyrics based on submitted lyrics." As the evidence cited does not support that statement, speaks for itself, and generally includes only short excerpts of lyrics. These records were generally created before Anthropic augmented its guardrails intended to prevent reproduction of copyrighted materials and specifically the works in suit stipulated with Publishers to maintain their efficacy; there is no evidence the lyrics in the cited documents are accurate. Pls.' Ex. 64 (Anthropic_0000459249) at -271; Pls.' Ex. 65 (Anthropic_0000557304) at -328; Pls.' Ex. 66 (Anthropic_0000459543) at -742; Pls.' Ex. 67 (Anthropic_0000557359), -612–14; Pls.' Ex. 79 (Anthropic_0000570595) at -596; Pls.' Ex. 85 (Dep. Ex. 215, Anthropic_0000524147) at -147; Pls.' Ex. 86 (Anthropic_0000431481); Pls.' Ex. 87 (Anthropic_0000430104); Pls.' Ex. 88 (Anthropic_0000430107); Pls.' Ex. 89 (Anthropic_0000434264); Pls.' Ex. 25 (Def.'s Am. Suppl. Resp. to Pls.' Interrog. No. 12, Nov. 12, 2025) at 11-14; Dkt. 291.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	Publishers' Additional Fact #43: Anthropic employees have prompted Claude to modify or combine existing song lyrics. **Pls.' Ex. 59** (Anthropic_0000016025) at -026-027; **Pls.' Ex. 92** (Anthropic_0000434832).	Undisputed that in Pls.' Ex. 59, a user asks Claude to "[w]rite a coherent poem made up of fragments of other poems" and then "[d]o the same with lyrics from Beatles, Bob Dylan, and other classics from the 60s/70s" and "with 90s and 00s hip hop." The outputs are not in the cited document and do not identify any work in suit. Undisputed that in Pls.' Ex. 92, the user writes, "I want new lyrics to we didn't start the fire. It's called we didn't start agi." Claude responds with lyrics to an original song, writing: "Here is an attempt at new lyrics for 'We Didn't Start AGI' chronicling the history of technology, while being careful not to reproduce any copyrighted material." Pls.' Ex. 59 (Anthropic_0000016025) at -035; Pls.' Ex. 92 (Anthropic_0000434832) at -832.
	Publishers' Additional Fact #44: Anthropic employees have prompted Claude with requests to produce AI-generated lyrics on given topics and in the style of existing songs or songwriters. **Pls.' Ex. 10** (Daniela Amodei Dep. Tr.) at 202:20–203:14; **Pls.' Ex. 56** (Anthropic_0000554366) at -388; **Pls. Ex. 62** (Anthropic_0000527903) at -913; **Pls.' Ex. 68** (Anthropic_0000557015) at -027; **Pls.' Ex. 69** (Anthropic_0000241820) at -820; **Pls.' Ex. 70** (Anthropic_0000459487) at -496; **Pls.' Ex. 71** (Dep. Ex. 244, Anthropic_0000554769) at -845-47; **Pls.' Ex. 72** (Anthropic_0000554405) at -413; **Pls.' Ex. 90**	Undisputed that each cited record reflects requests for new, original lyrics and none show Claude reproducing an existing work. Undisputed that, in Pls.' Ex. 10, Ms. Daniela Amodei acknowledged that an Anthropic employee sang a new and original "song written by CLAUDE." Pls.' Ex. 10 (Daniela Amodei Dep.) at 202:11–203:14. Undisputed that, in Pls.' Ex. 56, Claude outputs lyrics to the song "Wernher von Braun" by Tom Lehrer and later provides an "adapted version of the lyrics on the topic of artificial intelligence." "Wernher von Braun" is not a works in suit. Disputed that this record shows that "employees have prompted" because the record does not contain the user's prompts of Claude.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	(Anthropic_0000430292); **Pls.' Ex. 91** (Anthropic_0000431502); **Pls.' Ex. 93** (Anthropic_0000432472); **Pls.' Ex. 109** (Anthropic_0000432334) at -334-35; **Pls.' Ex. 110** (Dep. Ex. 174, Anthropic_0000570711); **Pls.' Ex. 111** (Dep. Ex. 242, Anthropic_0000570586).	Pls.' Ex. 56 (Anthropic_0000554366) at -388. Undisputed that, in Pls.' Ex. 62 does Claude provides an "attempt at a song called 'Tokenizer' to the tune of 'Womanizer' by Britney Spears." "Womanizer" is not a work in suit. Disputed that this record shows that "employees have prompted" because the record does not contain the user's prompts of Claude. Pls.' Ex. 62 (Anthropic_0000527903) at -913. Undisputed that, in Pls.' Ex. 68 Claude shares an "attempt at rewriting the lyrics of The Smiths' 'Frankly Mr. Shankly' to be about Twitter and science instead of the music industry." "Frankly Mr. Shankly" is not a work in suit. Disputed that this record shows that "employees have prompted" because the record does not contain the user's prompts of Claude. Pls.' Ex. 68 (Anthropic_0000557015) at -027-028. Undisputed that in Pls.' Ex. 69, the user requests "'oops i did it again' parody lyrics but it's about another android 0-day vulnerability patch." Claude outputs original lyrics. "Oops! I Did It Again" is not a work in suit. Pls.' Ex. 69 (Anthropic_0000241820) at -820-828. Undisputed that, in Pls.' Ex. 70 Claude shares an "attempt at continuing the parody lyrics," which include "Grilling me softly... with his tongs." "Killing Me Softly With His Song" is not a work in suit. Disputed

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		that this record shows that "employees have prompted" because the record does not contain the user's prompts of Claude. Pls.' Ex. 70 (Anthropic_0000459487) at -496. Undisputed that, in Pls.' Ex. 71 Claude generates new and original lyrics criticizing Donald Trump, Joe Biden, Barack Obama, and others. Disputed that this record shows that "employees have prompted" because the record does not contain the user's prompts of Claude. Pls.' Ex. 71 (Dep. Ex. 244, Anthropic_0000554769) at -845-847. Undisputed that, in Pls.' Ex. 72 Claude outputs an original "song about cloud computing" called "The Cloud," as well as a parody of Edgar Allen Poe's poem, "The Raven." Disputed that this record shows that "employees have prompted" because the record does not contain the user's prompts of Claude. Pls.' Ex. 72 (Anthropic_0000554405) at -413. Undisputed that in Pls.' Ex. 90, the user asks Claude to "[w]rite a song that combines the themes of climate change, ai existential risk, spike ball, nuclear wear, 10 minute planking, losing at card games, toddlers" to "the tune of" Elton John's "Tiny Dancer." Pls.' Ex. 90 (Anthropic_0000430292). Publishers also cannot demonstrate the scope of rights they control for Tiny Dancer, a work in suit, because they did not produce a true and correct

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		copy of the deposit made with the United States Copyright Office in connection with any application for registration for this work. Pls.' Ex. 64 (Anthropic_0000459249) at -271; Ex. 29 (Pls.' Supp. Resp. & Obj. to Def.'s 3rd Set of Req. For Admis. (Oct. 9, 2025)) at No. 929, p. 416. Undisputed that in Pls.' Ex. 91, the user requests, "[p]lease write a song about the ML concept of 'token healing' in the style of … 'Sexual Healing,'" and Claude responds with an "attempt at a lighthearted song about token healing." "Sexual Healing" is not a work in suit. Pls.' Ex. 91 (Anthropic_0000431502). Undisputed that in Pls.' Ex. 93, the user requests, "[p]lease write a song called '[I]'m just a lawyer' in the style of I'm just a girl by no doubt," and Claude responds, "Unfortunately, I cannot reproduce full copyrighted song lyrics," then provides a "summary of what an original parody song called 'I'm Just a Lawyer' … could cover." "I'm Just a Girl" is not a work in suit. Pls.' Ex. 93 (Anthropic_0000432472). Undisputed that, in Pls.' Ex. 110, Ms. Daniela Amodei asked Claude, "Can you write lyrics about a pig to the childrens song of baa baa black sheep for me to sing to my toddler?" Pls.' Ex. 110 does not include a request for lyrics written in the style of an existing copyrighted song or songwriter.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		Pls.' Ex. 110 (Dep. Ex. 174, Anthropic_0000570711); *see also* Pls.' Ex. 10 (Daniela Amodei Dep.) at 186:11-16. Disputed that Pls.' Ex. 109 includes a request for "AI-generated lyrics on given topics and in the style of existing songs or songwriters." The user requests that Claude "come up with a jingle [the user] can sing" to their dog. Claude outputs an original song, and states, "[f]eel free to sing this jingle to her in whatever tune you'd like." When the user asks, "[w]hat tune would you recommend?" Claude suggests "some fun options you could try singing it to," including "If You're Happy and You Know It," Take Me Out to the Ball Game," "Clementine," and "She'll Be Comin Around the Mountain." None are works in suit. Pls.' Ex. 109 (Anthropic_0000432334) at -334-335. Disputed that Pls.' Ex. 111 includes a request for "AI-generated lyrics on given topics and in the style of existing songs or songwriters." The user requests, "[w]rite a cynical dark but hilarious Rock song with the word weep." Claude outputs lyrics to a "satirical rock song that combines dark humor with commentary on consumer culture, debt, and the modern workforce." Pls.' Ex. 111 (Dep. Ex. 242, Anthropic_0000570586), at -586-587.
	Publishers' Additional Fact #45: Anthropic employee Ben Kuhn prompted Claude for help "putting together a book of song lyrics" by accessing and downloading	Undisputed that in Pls.' Ex. 162, the user states, "I'm putting together a book of song lyrics. How should I format/typeset the title, byline, verses, and chorus?" and Claude responds

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	copyrighted sheet music from a composer's website, and caused Claude to take steps to circumvent websites' "access restrictions" and "security measures" protecting against such unauthorized downloads. **Pls.' Opp'n/Reply Ex. 162** (Anthropic_0000570572)**; Pls.' Opp'n/Reply Ex. 163** (Anthropic_0000524116).	with "some standard formatting guidelines for song lyrics in a book." Pls.' Opp'n/Reply Ex. 162 (Anthropic_0000570572). Disputed that Claude actually circumvented "access restrictions" and "security measures" to "download" a "book of song lyrics." Pls.' Ex. 163 ends with Claude's output: "I apologize, but I've encountered a significant limitation: the website has security measures (403 Forbidden errors) that prevent automated downloading of the PDFs. Even using various techniques (wget, curl, Python requests with browser headers), I cannot bypass these protections … Would you like to try downloading them manually and uploading them to me? Or if you have another way to provide access to these files, let me know!" Pls.' Opp'n/Reply Ex. 163 (Anthropic_0000524116), at -144.
	Publishers' Additional Fact #46: Claude has generated output copying Publishers' lyrics even when users do not prompt Claude for those lyrics. **Pls.' Ex. 19** (Compilation of third-party prompts and outputs) at Anthropic_0000027609 (October 9, 2023) (generating the lyrics to "Just Like Heaven" in response to the prompt "finish this sentence: what do birds suddenly appear. . . what if I said the sentence was part of a song?").	Disputed that Claude generated lyrics even when users do not prompt Claude for those lyrics as the cited exhibit contains a prompt for lyrics. The user wrote, "finish this sentence: what do birds suddenly appear…" and Claude responded, "Here is one way to finish that sentence: What do birds suddenly appear when the sky is clear and blue? The user then wrote, "what if I said the sentence was ***part of a song***?" and Claude responded, "Okay, if that sentence is part of a song lyric, then I think I know the full line: 'What do birds suddenly appear every time you are near?'"

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		Pls.' Ex. 19 (Compilation of third-party prompts and outputs) at Anthropic_0000027609 (p. 183) (emphasis added). Disputed as to the characterization of Pls.' Ex. 19 to the extent that "users" refers to third party users. Publishers have not shown that the prompts in this exhibit were made by bona fide third parties. *See* Ex. 115 (Dennis Dep.) at 242:6-10; Ant. Reply Ex. 161 (Dennis Opening Report) ¶¶ 56-57. Further disputed because nearly all prompts in the cited exhibit were caused by users before Anthropic augmented its guardrails intended to prevent the generation of outputs that may infringe copyrights and specifically the works in suit and stipulated with Publishers to maintain their efficacy. Pls.' Ex. 25 (Def.'s Am. Suppl. Resp. to Pls.' Interrog. No. 12, Nov. 12, 2025) at 11-14; Dkt. 291.
	Publishers' Additional Fact #47: Claude is capable of generating output containing copies of song lyrics because those same song lyrics are included in Claude's training data. **Burrell Decl., Ex. 10** (Zhao Report) at ¶¶ 69–70, 92; **Burrell Decl., Ex. 18** (Zhao Rebuttal) at ¶¶ 109–10, 112; **Pls.' Opp'n/Reply Ex. 129** (Brown Dep. Tr.) at 84:23–85:1; *see also* **Pls.' Ex. 10** (Daniela Amodei Dep.) at 189:20–190:1.	Undisputed that Claude's training data included publicly available webpages containing lyrics. Disputed that Claude retrieves or stores copies of lyrics in its training data. Anthropic's expert, Dr. Jeffrey Bilmes, reported: "Anthropic uses training data to train its models for the purpose of mapping out statistical relationships between text fragments so that the LLM models can generate new texts. Those statistical relationships are unrecognizable transformations of the massive

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		quantity of training data used to train a model such as Claude and are not derivative works of Publishers' lyrics." "LLMs do not store or retrieve passages of text verbatim. Instead, they generate new and different outputs, probabilistically, based on general statistical patterns and relationships learned during training that are represented in the model's weights. Because their outputs are drawn from probability distributions, LLMs can produce different results when generating outputs in response to the same prompt, particularly when the final text is produced by stochastically (i.e., randomly) sampling from those distributions, which is by far the most common approach to LLM generation." Ex. 2 (Bilmes Report) ¶¶ 36, 98. Publishers' expert, Dr. Ben Zhao, testified that Claude's "training corpus is decoupled from the model once the model is trained"; that Claude does not "stor[e] an exact copy of that text"; and that "the model does not scan for the text from its training corpus." Ex. 139 (Zhao Dep.) at 177:2-4, 178:6-18, 179:9-10.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
Issue 3: The Second Factor Favors Fair Use Because Song Lyrics are Freely and Publicly Available	Fact 51: Musical compositions and lyrics can be creative, expressive works. Ex. 137 (Stokes Dep.) at 68:9-16, 224:21-24	Undisputed.
	Fact 52: In response to the question "Consumers expect to be able to get lyrics online for free, right?" 30(b)(6) Representative for UMPG, Mr. David Kokakis, testified "I know that's how we see most lyrics-based digital services offering lyrics to consumers." Ex. 121 (Kokakis Dep.) at 250:8-253:13	Disputed. Mr. Kokakis testified in full, "I don't know what the consumers expect. I know that's how we see most lyrics-based digital services offering lyrics to consumers." **Pls.' Opp'n/Reply Ex. 143** (Kokakis Dep.) at 251:22–252:3.
	Fact 53: In response to the question "Is there any website that you have gone to to view the lyrics for particular songs?" 30(b)(6) Representative for Concord, Mr. Duff Berschback, testified "I'll sometimes use Genius and sometimes do just a simple Google search, both of which are -- have been licensed by Concord." And in response to the question "So with respect to genius.com, for example, anyone could go online and view the lyrics to any of Concord's works that are included on genius.com?" 30(b)(6) Representative for Concord, Mr. Duff Berschback, testified "Yes." Ex. 104 (Berschback Dep.) at 122:24-123:16	Undisputed.

Publishers' Additional Material Facts on This Issue	Anthropic's Response & Supporting Evidence
Publishers' Additional Fact #48: Publishers' licensees and sub-licensees, including lyric display websites such as Genius.com, pay	Undisputed that Publishers' licenses with lyric display websites include contractual provisions relating to

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	monetary consideration for the rights to display or use Publishers' lyrics. **Smith Decl. ¶¶** 13–15, 109; **Pls.' Ex. 101** (UMPG_0000018093, Concord_0000000168, ABKCO_0000000052, CCMG_0000000209) (LyricFind licenses); **Pls.' Ex. 102** (UMPG_0000000335, ABKCO_0000000045, CCMG_0000000237, Concord_0000000212) (MusixMatch licenses); **Pls.' Ex. 103** (UMPG_0000000388, UMPG_0000023282, UMPG_0000023090, UMPG_0000023257, UMPG_0000000275, UMPG_0000000366, ABKCO_0000000414, CCMG_0000000869, Concord_0000000739, UMPG_0000000532) (compilation of additional lyrics display licenses); *see also* **Pls.' Opp'n/Reply Ex. 167** (UMPG_0000000311; UMPG_0000023040; UMPG_0000023125; ABKCO_0000000328; ABKCO_0000000025; ABKCO_0000001313; ABKCO_0000000304; ABKCO_0000000369; ABKCO_0000000398; CCMG_0000000662; CCMG_0000000259) (compilation of additional lyrics display licenses).	monetary consideration for the rights to display or use Publishers' lyrics.
	Publishers' Additional Fact #49: Many lyrics display websites, including Publishers' licensees, "tend to be ad-supported, which doesn't necessitate the user making a payment for access to lyrics, but	Undisputed that users do not "mak[e] a payment for access to lyrics" and that "the user doesn't have to pay for it." When asked whether "[c]onsumers expect to be able to get lyrics online for free, right," 30(b)(6)

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	the service itself picks up the cost out of the advertising revenue that's generated[,] [s]o the cost to consumer isn't a cash payment. The cost to the consumer is their attention," much like broadcast TV and radio "where there are interstitial commercials running throughout programming and the user doesn't have to pay for it." **Pls.' Opp'n/Reply Ex. 143** (Kokakis Dep. Tr.) at 250:25–251:20.	Representative for UMPG Mr. David Kokakis testified, "I don't know what consumers expect. I know that that's how we see most lyrics-based digital services offering lyrics to consumers." Ant. Reply Ex. 166 (Kokakis Dep.) at 250:25–252:3; *id.* at 250:8-253:13.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
Issue 4: The Third Factor Favors Fair Use Because Only the Portion of The Works Reasonably Necessary to Achieve a Transformative Purpose Is Used	Fact 54: Publishers admit that "the copyrighted musical compositions for which they own and control exclusive rights, including the Asserted Works, include music and lyrics." Ex. 27 (Pls.' Resp. & Obj. to Def.'s 2nd Set of Req. For Admis. (Nos. 16-24) (Sept. 2, 2025)) at 9	Undisputed.
	Fact 55: 30(b)(6) Representative for UMPG, Mr. David Kokakis, testified [redacted] Ex. 121 (Kokakis Dep.) at 85:9-87:14	Undisputed.
	Fact 56: When asked "ABKCO does not have separate copyrights to the lyrics for any of the songs identified in the ABKCO song list, right?" 30(b)(6) Representative for ABKCO, Ms. Alisa Coleman, answered "Correct." Ex. 110 (Coleman Dep.) at 49:22-50:6	Undisputed.
	Fact 57: When asked "Lyrics are not the entirety of any of the 40 Concord works in suit?" 30(b)(6) Representative for UMPG, Mr. Duff Berschback, testified "That's correct." Ex. 104 (Berschback Dep.) at 30:24-31:7	Undisputed.
	Fact 58: When asked [redacted] 30(b)(6) Representative for CCMG, Mr. Draughon, testified [redacted] He also testified	Undisputed.

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	[redacted] Ex. 117 (Draughon Dep.) at 50:13-17, 51:10-18	
	Fact 59: Publishers' expert, Dr. Shawn Shan, reported "to acquire these sufficiently large training corpora, developers often turn to a massive source of easily accessible text: the Internet." He testified "The largest portion of [Claude's] training data consists of 'Common Crawl' data. Common Crawl is a collection of website crawls from 2008 to the present, including raw web pages as well as corresponding metadata and text extractions." Ex. 7 (Shan Report) at ¶¶ 22, 45(i)	Undisputed, except that Dr. Shan did not "testif[y]" as quoted. The quoted language appears in Dr. Shan's opening expert report. **Burrell Decl., Ex. 7** (Shan Report) at ¶ 45(i).
	Fact 60: Anthropic's expert, Dr. Jeffery Bilmes, reported that publicly available information indicates Anthropic's Claude models were trained in part on data sourced from Common Crawl or by directly crawling the web. Ex. 2 (Bilmes Report) at ¶¶ 102-103, 107	Undisputed.
	Fact 61: Publishers' Expert, Dr. Ben Zhao, reported "Anthropic has copied the lyrics to a huge number of Publishers' musical compositions and used those copies as input to train Claude." Ex. 10 (Zhao Report) at ¶ 10	Undisputed.
	Fact 62: Publishers' Expert, Dr. Shawn Shan, reported "Anthropic trains its Claude AI models using certain datasets—including Common Crawl, OpenWebText2, The Pile, YouTube Subtitles, and	Undisputed.

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	Reddit data—that contain Publishers' lyrics." Ex. 7 (Shan Report) at ¶ 44	
	Fact 63: In response to the question [redacted] (b)(6) Representative for CCMG, Mr. Draughon, testified [redacted] Ex. 117 (Draughon Dep.) at 136:21-137:2	Undisputed that Mr. Draughon testified as quoted. Anthropic relies on this fact for the truth of the matter asserted that Anthropic has not copied, used, or otherwise accessed the entirety of the musical composition, including the lyrics and the music, for any Capitol CMG work. Publishers dispute the truth of the matter asserted that Anthropic has not copied, used, or otherwise accessed the entirety of the musical composition, including the lyrics and the music, for any Capitol CMG work. *See* **Pls.' Opp'n/Reply Ex. 169** (ANTH_DATABASE_0000005467 ("Trending Worship Songs" with sheet music and lyrics to CCMG Works in Suit for Piano/Vocal/Guitar); ANTH_DATABASE_0000005062 ("Top Hits of 2018" with sheet music and lyrics to CCMG Works in Suit for Piano/Vocal/Guitar); ANTH_DATABASE_0000005478 ("Top Hits of 2018" with sheet music and lyrics to CCMG Works in Suit for Easy Piano)).
	Fact 64: Audio files and data are not supported by Claude models and thus are not incorporated in the models' training data. Ex. 11 (Bilmes Rebuttal) at ¶ 239	Undisputed.
	Fact 65: Publishers' expert, Dr. Ben Zhao, reported "[r]egardless of their specific task or architecture, generative AI systems are trained on very large datasets of content."	Undisputed, except that Ex. 10 (Zhao Report), not Ex. 18, contains the quoted language.

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	He reported that "[t]he larger the corpus on which a model is trained, the more effective and reliable its results will be as a general matter." Ex. 18 (Zhao Report) at ¶¶ 22, 23	**Burrell Decl., Ex. 10** (Zhao Report) ¶ 23.
	Fact 66: Anthropic co-founder and Chief Science Officer, Dr. Jared Kaplan, attested that "[t]rillions of tokens are required to properly train an LLM like Claude and to ultimately allow the LLM to recognize patterns in language to any useful degree." Kaplan PI Decl. ¶ 22	Undisputed that Dr. Kaplan attested to the quoted language. Anthropic relies on this fact for the truth of the matter asserted by the quoted language. Publishers dispute the truth of the matter asserted that "trillions of tokens are required to properly train an LLM like Claude and to allow the LLM to recognize patterns in language to any useful degree." For example, GPT-3, a "general-purpose LLM" like Claude, was trained on 300 billion tokens. **Pls.' Opp'n/Reply Ex. 172** (Dep. Ex. 202, Tom B. Brown et al., *Language Models Are Few-Shot Learners*, ARXIV:2005.14165v4, 8–9 (July 22, 2020); **Pls.' Opp'n/Reply Ex. 137** (Nagaraj Dep. Tr.) at 36:23–37:9, 42:2-8, 44:4–45:16; *see also* **Kaplan Decl.** ¶¶ 25–26 (recognizing that LLMs can be trained on smaller corpora of data and remain effective for at least certain tasks); **Pls.' Ex. 6** (Kaplan Dep. Tr.) at 138:10-14 (testifying "that's not true" in response to the question: "Anthropic wants large volumes of training data; is that true?"); **Pls.' Opp'n/Reply Ex. 129** (Brown Dep. Tr.) at 92:2-7 (testifying "the quality of the data is much more important than the overall quantity of the data"); **Burrell Decl., Ex. 18** (Zhao Rebuttal) at ¶ 23; **Pls.' Opp'n/Reply Ex. 145** (Shan Dep. Tr.) at 105:7-10.
	Fact 67: Publishers' expert, Dr. Shawn Shan, testified "[Y]es," that	Undisputed.

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	it was accurate that "LLM developers like Anthropic seek to train LLMs like Claude on a vast training corpus of human written texts often comprising of terabytes or petabytes of data and hundreds of billions or trillions of tokens." Ex. 131 (Shan Dep.) at 125:10-128:6	
	Fact 68: Publishers' expert, Dr. Ben Zhao, agreed that "song lyrics are available on the internet through a number of different sites," testifying that "they are available in different forms, different places." Ex. 139 (Zhao Dep.) at 276:19-25	Disputed. The quoted language does not accurately reflect Dr. Zhao's testimony. In response to the question, "[Y]ou understand that song lyrics are widely available on the internet through a number of different sites, right?", Dr. Zhao testified in full: "I understand that they are available in different forms, different places. I also understand that many of these places have actual licensed access to them and agreements with some of the plaintiffs and other music and music industry companies." **Burrell Decl., Ex. 139** (Zhao Dep.) at 276:19-25.
	Fact 69: Publishers' expert, Dr. Shawn Shan, reported that to "acquire these sufficiently large training corpora, developers often to turn to a massive source of easily accessible text: the Internet. Developers typically source large portions of their training corpora by scraping text from webpages, a process by which bots 'crawl' through the internet and copy the contents of webpages it encounters. Developers may also acquire datasets constructed by third parties, such as 'Common Crawl,'" 'The Pile,' and 'OpenWebText2,' which are	Undisputed.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	themselves largely comprised of petabytes of content scraped from the Internet." Dr. Shan also reported that "Web crawling is an automated process by which bots copy the contents of webpages they come across on the Internet." Ex. 7 (Shan Report) at ¶¶ 22-23	
	Fact 70: Anthropic's expert, Dr. Jeff Bilmes, reported "Without large text collections like Common Crawl and other large-scale web crawling efforts, it would be very difficult to obtain sufficient data for the types of advanced deep LLM training that has made general purpose LLMs like Claude possible." And that "Developers typically meet their data requirements by gathering training data from several sources, including large publicly available text corpora like Common Crawl and OpenWebText. Common Crawl and OpenWebText contain textual information extracted from millions of webpages over time." Ex. 2 (Bilmes Report) at ¶¶ 100, 102	Undisputed that Dr. Bilmes' report contains the quoted language.
	Fact 71: Claude Sonnet 3.0 was trained on [redacted] exceeds more than twice [redacted] of the text of all the books in the world. Ex. 6 (Nagaraj Report) at ¶¶ 36-37; Kaplan PI Decl. ¶¶ 29-30; Ex. 56 (Anthropic_0000231226)	Disputed. Claude Sonnet 3.5 (released on June 20, 2024), not Claude Sonnet 3.0, was trained on [redacted]. Claude Sonnet 3.0 (released on March 4, 2024) was trained on [redacted] **Burrell Decl., Ex. 6** (Nagaraj Report) at ¶¶ 36–37; Kaplan PI Decl. ¶ 29; **Pls.' Opp'n/Reply Ex. 173** (Anthropic_0000592435) (ANTHROPIC, *Introducing the next generation of Claude*); **Pls.'**

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
		Opp'n/Reply Ex. 174 (Anthropic_0000592449) (ANTHROPIC, *Claude 3.5 Sonnet*). Further disputed because Ex. 56 does not support the asserted fact, and only reflects Anthropic's [redacted] ***words*** in all the books in the world, not tokens. *See* **Burrell Decl., Ex. 56** (Anthropic_0000231226).
	Fact 72: Publishers' expert, Dr. Shawn Shan, testified that "the data collection happens across many years, including sourcing from Common Crawl" and "data sharks." He also testified "based on my inspection, there are large amount of work in suit as part of internet data." Ex. 131 (Shan Dep.) at 232:5-233:4, 343:2-22, 344:1-16	Undisputed that Dr. Shan testified, "based on my inspection, there are large amount of work in suit as part of internet data." Otherwise, disputed. The quoted language does not accurately reflect Dr. Shan's testimony. Dr. Shan testified: "So based on my analysis of how Anthropic collected data, the data collection happens across many years, including sourcing from Common Crawl and ***the, you know, data sharks, which was Common Crawl***—that Anthropic leveraged trace back to the dates, I believe, 2014." **Burrell Decl., Ex. 131** (Shan Dep.) at 232:20-25 (emphasis added).
	Fact 73: Publishers' expert, Dr. Ben Zhao, testified that "Anthropic has gotten training data from a number of sources" such as "web archives, YouTube video transcriptions, Common Crawl in the past." When asked whether he "understand[s] that song lyrics are widely available on the internet through a number of different sites," Dr. Zhao responded that he understands that song lyrics "are available in different forms, different places." Publishers' expert, Dr. Shawn Shan, reported that "Anthropic trains its Claude AI models using certain datasets—including Common Crawl,	Undisputed that Dr. Shan's report quotes the creators of the Pile dataset in stating that OpenWebText2 is "a generalized web scrape dataset . . . includ[ing] more recent content from Reddit submissions up until 2020." Also undisputed that Dr. Shan's report states that "Anthropic trains its Claude AI models using certain datasets—including Common Crawl, OpenWebText2, The Pile, YouTube Subtitles, and Reddit data—that contain Publishers' lyrics" and that Common Crawl "represents a massive indiscriminate scrape of the public web."

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	OpenWebText2, The Pile, YouTube Subtitles, and Reddit data—that contain Publishers' lyrics" that CommonCrawl "represents a massive indiscriminate scrape of the public web" and that "OpenWebText2" is "a generalized web scrape dataset." Ex. 139 (Zhao Dep.) at 274:20-275:7, 276:19-277:4, 279:8-280:3, 282:2-21; Ex. 7 (Shan Report) ¶¶ 44-45	Otherwise, disputed. The quoted testimony does not accurately reflect Dr. Zhao's testimony. When asked, "Are you aware of any evidence that Anthropic ever scraped lyrics for any songs, or for the works in suit in particular, without respecting robots.txt?", Dr. Zhao responded in full: "So let me unpack that question for a second. Am I aware—do I have specific knowledge that Anthropic explicitly ignored robots.txt in scraping lyric content for its training data? The answer to that question is no. However, Anthropic has gotten training data from a number of sources, not all of which have respected robots.txt; so web archives, YouTube video transcriptions, Common Crawl in the past. So all of these things are things that do not require Anthropic to interact with robots.txt. Even though someone else may have gotten data on their behalf or relayed them—data to them through some means and if they did not respect robots.txt to get that data, Anthropic would not have necessitated—would not have been necessary for Anthropic to personally and directly ignore robots.txt, and, yet, they would still have been privy to data that was obtained through that process." **Burrell Decl., Ex. 139** (Zhao Dep.) at 274:20–275:23. In response to the question, "[Y]ou understand that song lyrics are widely available on the internet through a number of different sites, right?", Dr. Zhao testified in full: "I understand that they are available in different forms, different places. I also understand that many of these places have actual licensed access to them

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
		and agreements with some of the plaintiffs and other music and music industry companies." **Burrell Decl., Ex. 139** (Zhao Dep.) at 276:19–277:4.
	Fact 74: Anthropic's expert, Dr. Jeffrey Bilmes, reported that Anthropic's general-purpose web crawler is designed to respect internet protocols that tell web crawlers not to crawl or scrape those web pages. Dr. Bilmes analyzed Anthropic's web crawling data and found that Anthropic collected lyrics to the works-in-suit from [redacted] million public web pages and adhered to the internet protocol rules in place at the time those web pages were crawled. Ex. 2 (Bilmes Report) at § VII.A	Undisputed that Dr. Bilmes' report states that "Anthropic's general-purpose web crawler is designed to respect internet protocols that tell crawlers not to crawl or scrape pages," and that Dr. Bilmes' report states that he analyzed Anthropic's web crawling data as asserted. Publishers dispute that Dr. Bilmes' analysis shows Anthropic "adhered to the internet protocol rules in place at the time" because Dr. Bilmes improperly includes websites without robots.txt files or instructions within his [redacted] million count, and Anthropic may have impermissibly crawled over [redacted] that prohibited or blocked Anthropic's crawler. *See* **Burrell Decl., Ex. 2** (Bilmes Report) at ¶ 306 & n.580; **Burrell Decl., Ex. 16** (Shan Rebuttal) at ¶¶ 18–20; **Burrell Decl., Ex. 18** (Zhao Rebuttal) at ¶¶ 73–84. Further disputed because the cited evidence (Def.'s Ex. 2) is subject to a pending motion to exclude. *See* ECF No. 642.

Publishers' Additional Material Facts on This Issue	Anthropic's Response & Supporting Evidence
Publishers' Additional Fact #50: Publishers frequently license just the lyrics to their copyrighted musical compositions to third parties. *See, e.g.*, **Pls.' Opp'n/Reply Ex. 161** (Pls.' Supp. Resp. to Def.'s	Undisputed that Publishers have licenses with certain licensees for the display of lyrics to musical compositions. Otherwise disputed that Publishers enter into licenses for just the lyrics

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	Interrog. No. 25, Sept. 18, 2025) at 8; **Pls.' Opp'n/Reply Ex. 140** (Berschback Dep. Tr.) at 127:3-15, 193:10-23; **Pls.' Opp'n/Reply Ex. 141** (Coleman Dep. Tr.) at 132:21–133:9, 133:24–135:23, 140:11–141:9, 144:8–145:2; **Pls.' Opp'n/Reply Ex. 142** (Draughon Dep. Tr.) at 50:22–51:3,107:22–108:4, 109:6-9, 198:22–199:16, 202:17–205:4, 212:7-25; **Pls.' Opp'n/Reply Ex. 143** (Kokakis Dep.) 124:12–125:1, 125:19–126:9, 236:4–238:1; *see also* Decl. of Duff Berschback ISO Pls.' Mot. for Summ. J. ("**Berschback Decl.**") (March 20, 2026), ¶¶ 42–45, ECF No. 598; Decl. of Alisa Coleman ISO Pls.' Mot. for Summ. J. ("**Coleman Decl.**") (March 20, 2026), ¶¶ 15, 17–21, ECF No. 596; Decl. of Kent Draughon ISO Pls.' Mot. for Summ. J. ("**Draughon Decl.**") (March 20, 2026), ¶¶ 32–34, ECF No. 599; Decl. of David Kokakis ISO Pls.' Mot. for Summ. J. ("**Kokakis Decl.**") (March 22, 2026), ¶¶ 146, 148–51, ECF No. 610.	"frequently" outside of lyric display aggregating websites. Pls' Opp'n/Reply Ex. 161 discloses only licenses to "URL domains and software applications" that "have executed licenses with one or more Publishers to use or display the lyrics" without regard to the frequency of those licenses. Pls.' Opp'n/Reply Ex. 161 (Pls.' Supp. Resp. to Def.'s Interrog. No. 25, Sept. 18, 2025) at 8. Mr. Berschback testified that "lyric aggregators, such as Musixmatch and Lyricfind are – they license our lyrics, and then they – in the case of lyrics find, they have their own consumer destination site." Pls.' Opp'n/Reply Ex. 140 (Berschback Dep.) at 127:3-15. Ms. Coleman's testimony regards licenses for "a book" and with "Genius Media Group" with "Musixmatch" and with "LyricFind." Pls.' Opp'n/Reply Ex. 141 (Coleman Dep.) at 132:21–133:9, 133:24–135:23, 140:11–141:9, 144:8–145:2. Mr. Draughon's testimony is about licenses with "LyricFind or Musixmatch" and "PraiseCharts" Pls.' Opp'n/Reply Ex. 142 (Draughon Dep.) at 50:22–51:3,107:22–108:4, 109:6-9, 198:22–199:16, 202:17–205:4, 212:7-25. Mr. Kokakis testified to licenses with "LyricFind" and "Genius."

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		Pls.' Opp'n/Reply Ex. 143 (Kokakis Dep.) at 124:12–125:1, 125:19–126:9, 236:4–238:1
	Publishers' Additional Fact #51: Claude was designed and trained to be a text-only LLM capable of "generat[ing] high quality text." **Pls.' Ex. 21** (Def.'s Am. Resp. to Pls.' RFAs, No. 80) at 83; **Pls.' Opp'n/Reply Ex. 13** (Dario Amodei Dep. Tr.) at 102:7-24; **Burrell Decl. Ex. 15** (Nagaraj Rebuttal) at ¶ 38.	Undisputed that Anthropic "cleans and processes the data used to train its models to remove data that would detract from the goal of teaching Claude to generate high quality text." Further undisputed that Claude "doesn't generate audio and video" and that Anthropic is "generally not aiming to produce content in creative industries." Further undisputed that "[a]s a text-based LLM, Claude does not have the capacity to generate full songs with lyrics *and* music." Pls.' Ex. 21 (Def.'s Am. Resp. to Pls.' RFAs, No. 80) at 83; Pls.' Ex. 13 (Dario Amodei Dep.) at 102:7-24; Ex. 15 (Nagaraj Rebuttal) at ¶ 38.
	Publishers' Additional Fact #52: Lyrics are considered an integral part of musical compositions. **Burrell Decl., Ex. 104** (Berschback Dep.) at 30:24–31:3.	Undisputed that lyrics can be part of a musical composition. Mr. Berschback also testified, "That's correct," in response to the question "Lyrics are not the entirety of any of the 40 Concord works in suit." Ex. 104 (Berschback Dep.) at 30:24–31:7. Further undisputed that none of the 499 works in suit is comprised only of lyrics: Publishers admit that "the copyrighted musical composition for which they own and control exclusive rights, including the Asserted Works, include music and lyrics." Ex. 27 (Pls.' Resp. & Obj. to Def.'s 2nd Set of Req. For Admis. (Nos. 16-24) (Sept. 2, 2025)) at 9; Ex. 121 (Kokakis Dep.) at 85:9-87:14; Ex. 110

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		(Coleman Dep.) at 49:22-50:6; Ex. 117 (Draughon Dep.) at 50:13-17, 51:10-18. "[L]yrics are only a part of the entire musical composition, if they are present at all." And while "[l]yrics can be a meaningful part of a composition [they] derive much of their value and expression from their setting within the rest of the musical composition and take on particular meaning when paired with melody, harmony, instrumentation, and other compositional elements." Ex. 9 (Stokes Report) ¶¶ 17(a), 34.
	Publishers' Additional Fact #53: Anthropic's stated policy is to license training data [redacted] from scraping or other Internet sources. **Pls.' Opp'n/Reply Ex. 148** (Dep. Ex. 35, Anthropic_0000229152) at -152–53; **Pls.' Opp'n/Reply Ex. 134** (Turvey 30(b)(1) Dep. Tr.) at 60:12–61:18; *see also* **Burrell Decl., Ex. 6** (Nagaraj Report) at ¶ 73.	Undisputed that the document cited contains the quoted language. Otherwise disputed for the truth of the matter asserted because in response to the question "[i]s this an official Anthropic policy document," about Pls.' Opp'n/Reply Ex. 148, 30(b)(6) Representative, Mr. Thomas Turvey, responded, "It is my document from my group … that's what it is," and that it was "a document that I wrote on my seventh day of employment in Anthropic and have updated since related to principles that I wanted … the company to take regarding how we would conduct data partnerships." Ant. Reply Ex. 167 (Turvey 30(b)(6) Dep.) at 108:23-110:15.
	Publishers' Additional Fact #54: A large volume of training data is not always desirable for LLM training; certain models perform better with a more limited training corpus. **Burrell Decl. Ex. 18** (Zhao Rebuttal Report) at ¶ 23; **Burrell**	Undisputed that certain models built for certain purposes may perform better with a more limited training corpus and that the quality of the data is an important factor in influencing model performance. Otherwise disputed as to the truth of the matter asserted because training

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	Decl. Ex. 2 (Bilmes Report) at ¶ 322; **Kaplan PI Decl.** ¶¶ 25–26, ECF No. 209 (recognizing that LLMs can be trained on smaller corpora of data and remain effective for at least certain tasks); *see also* **Pls.' Opp'n/Reply Ex. 128** (Kaplan Dep. Tr.) at 138:10-14 (testifying "that's not true" in response to the question: "Anthropic wants large volumes of training data; is that true?"); **Pls.' Opp'n/Reply Ex. 129** (Brown Dep. Tr.) at 92:2-7 (testifying "the quality of the data is much more important than the overall quantity of the data").	general-purpose LLMs, like Claude, requires large, diverse data. It is undisputed that Dr. Bilmes reported that "[h]eterogenous data sources such as Common Crawl, OpenWeb, and book data, have been shown to be most influential in producing strong downstream LLM performance…. [R]emoving large heterogenous datasets (specifically Common Crawl, OpenWeb, and book data) from the pre-training corpus had a greater impact on model performance than removing more narrow specialized domains." Ex. 2 (Bilmes Report) ¶ 321. It is also undisputed that "scaling laws mean that as model builders create larger models using more computation, they also need more data. In other words, better model training is a function of model size, dataset size, and computation amount, among other factors, all in conjunction." Language models trained on less data may still achieve impressive results "albeit only on a subset of tasks for which general purpose models," like Claude, "excel." The developers of the small language model Dr. Zhao discussed at ¶ 23 of Ex. 18 recognized that "LLMs like Claude are still the 'gold standard'" and are not "equivalent to LLMs trained at vastly greater scale," in terms of performance, reasoning, and complex task completion. Ex. 19 (Bilmes Reply) ¶¶ 49-51. It is further undisputed that "Future models will require even more data to

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		improve upon performance. [Anthropic] expect[s] that subsequent models will also require tens of trillions of tokens to continue to show performance gains." Kaplan PI Decl. ¶ 30, Dkt. 212-5. A heterogenous large corpus of data, including low quality or value data is necessary to increase model performance generally and Anthropic has "empirically validated" that "throw[ing] out every document that the [redacted]" "would likely significantly skew [its] data distribution in a way that damaged model performance." Ex. 124 (Mann 30(b)(1) Dep. Vol. 1) at 82:11-16.
	Publishers' Additional Fact #55: Large datasets can sometimes result in decreased model performance. **Burrell Decl. Ex. 2** (Bilmes Report) at ¶ 123.	Disputed insofar as Publishers mischaracterize the evidence cited and as to the truth of the matter asserted because training general-purpose LLMs, like Claude, requires large, diverse data. It is undisputed that "LLMs, by virtue of their scale, require massive amounts of training data to achieve improvements in model performance…. The ***original*** scaling laws [] revealed a point of diminishing returns: ***if a model size were held constant*** while the size of training datasets continued to grow, the model would eventually have reduced performance…. Later research found that **compute-optimal** LLMs—LLMs that reach the best performance with the computation budget available—required fewer parameters ***and far larger training datasets*** than had been used for the very large (by parameter-

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		size) LLMs trained in 2021 and 2022." Ex. 2 (Bilmes Report) ¶¶ 122-123, 165 (italics emphases added). "[S]caling laws mean that as model builders create larger models using more computation, they also need more data. In other words, better model training is a function of model size, dataset size, and computation amount, among other factors, all in conjunction." Ex. 19 (Bilmes Reply) ¶¶ 49-51. A heterogenous large corpus of data, including low quality or value data is necessary to increase model performance generally and Anthropic has "empirically validated" that "throw[ing] out every document that the [redacted]" "would likely significantly skew [its] data distribution in a way that damaged model performance." Ex. 124 (Mann 30(b)(1) Dep. Vol. 1) at 82:11-16.
	Publishers' Additional Fact #56: Anthropic hired a contractor, Gyula Lakatos, to build its own web crawling and scraping infrastructure in 2023. **Pls.' Opp'n/Reply Ex. 124** (Mann 30(b)(6) Dep. Tr.) at 62:9-12.	Undisputed that Gyula Lakatos "was a contractor that Anthropic hired to help us build some web crawling and scraping infrastructure, originally, as I mentioned, focused only on these documents and not general web pages." Pls.' Opp'n/Reply Ex. 124 (Mann 30(b)(6) Dep.) at 62:9-14.
	Publishers' Additional Fact #57: Anthropic's first web crawler [redacted]	Undisputed that when Anthropic began its crawling activities [redacted], it [redacted] and

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	Pls.' Ex. 4 (Mann 30(b)(6) Dep. Tr.) at 56:18–57:21; **Pls.' Ex. 24** (Dep. Ex. 129, Anthropic_0000343228); **Pls.' Opp'n/Reply Ex. 151** (Dep. Ex. 130, Anthropic_0000342806); **Pls.' Opp'n/Reply Ex. 135** (Sangani Dep. Tr.) at 215:1–216:2 (discussing Dep. Ex. 130).	that "[t]here have been periods of time when we didn't [honor robots.txt instructions] due to bugs, and there have also been periods of time in the very early days when we were only crawling the specific documents types [] described [above], where, to the best of [Anthropic's] knowledge, it seems possible that we were not respecting robots.txt. But … [redacted] Pls.' Ex. 4 (Mann 30(b)(6) Dep.) at 56:18–57:21.
	Publishers' Additional Fact #58: Dr. Bilmes' analysis of Anthropic's general-purpose web crawler's adherence to internet protocols was limited to Anthropic's web crawler that launched in 2024 and the data scraped after this crawler was launched. **Burrell Decl., Ex. 2** (Bilmes Report) at ¶¶ 302–05; **Pls.' Ex. 4** (Mann 30(b)(6) Dep. Tr.) at 72:23–73:8.	Undisputed that Dr. Bilmes's analysis was limited to Anthropic's web crawler that launched in 2024. Anthropic's new [redacted] that uses [redacted] Pls.' Ex. 4 (Mann 30(b)(6) Dep.) at 71:18–73:8. Further undisputed that Dr. Bilmes analyzed a crawl dataset "which represent the collection of webpages Anthropic directly crawled between March 12, 2024, and January 31, 2025." Ex. 2 (Bilmes Report) ¶¶ 302–305.
	Publishers' Additional Fact #59: Dr. Bilmes' analysis of Anthropic's crawling activity demonstrated that for 1.88% of the URLs Anthropic crawled (over [redacted]),	Undisputed that 1.88% of the URLs Anthropic crawled, and analyzed by Dr. Bilmes, returned a "Disallow" directive for either Anthropic's crawler user-agent or all user agents and that approximately 75% of those

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	Anthropic's crawler proceeded to crawl despite URLs containing robots.txt directives that returned a "Disallow" signal. **Burrell Decl. Ex. 16** (Shan Rebuttal Report) at ¶ 18; **Burrell Decl., Ex. 2** (Bilmes Report) at ¶¶ 304–07.	URLs [redacted] came from the same domain, Lyricmeaning.com, which is no longer accessible and whose robots.txt file, beginning in 2022 and before Anthropic began crawling webpages, was unreachable due to action taken by the domain owner, making it impossible for any web crawler to determine the website's crawling directives. Undisputed that excluding those URLs demonstrates that Anthropic permissibly crawled 99.50% of URLs. Ex. 2 (Bilmes Report) at ¶¶ 304–307.
	Publishers' Additional Fact #60: Publishers' expert Dr. Shan conducted an independent empirical analysis of 10,000 unique web domains and determined that, under Dr. Bilmes' methodology, 94.35% of domains would be classified as permitting a crawl. **Burrell Decl. Ex. 16** (Shan Rebuttal Report) at ¶ 21 & nn.15–16.	Undisputed that Dr. Shan's report indicates he conducted an analysis of 10,000 unique domains extracted from the Tranco Top 1 Million list and that his "results closely match the characterization provided in the Bilmes Report." Otherwise disputed as to the truth of the matter asserted. Dr. Shan's analysis merely confirms that the robots.txt exclusion protocol is industry standard and that where user-agents are not disallowed from crawling, those web pages permit crawling. Ex. 16 (Shan Rebuttal Report) ¶ 21; Ex. 19 (Bilmes Reply Report) ¶ 73.
	Publishers' Additional Fact #61: Because the general architecture of the internet includes frequent absence of explicit technical barriers to crawling, compliance with such configurations does not reflect significant effort. **Burrell Decl. Ex. 16** (Shan Rebuttal Report) at ¶ 21 & nn.15–16.	Disputed. The robots.txt Robots Exclusion Protocol was defined in 1994 and has been published as an "Internet Standard" that is "interoperable and widely deployed" since 2022. Ex. 19 (Bilmes Reply Report) ¶ 73 & n.269. Publishers have directed licensees of their lyrics, in 2025, to include

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		robots.txt files that restrict [redacted] Ant. Reply Ex. 168 (Kokakis Dep. Ex. 25, UMPG_0000023303) at -305–307. A study conducted in 2025 by Dr. Zhao and Dr. Shan found that the Robots Exclusion Protocol "is now widely used to manage content access." And that Anthropic's robot.txt crawler's user-agents "are the two most-frequently restricted AI user agents" and that its user-agent "ClaudeBot" "respected the robots.txt file" in the study and found "Yes" ClaudeBot "Respect[s] [robot.txt] in practice." Ant. Reply Ex. 169 (Zhao Dep. Ex. 3, *Somesite I Used To Crawl*) at pp. 2, 4, 8, 10. Further disputed that compliance with robots.txt does not reflect significant effort because "[f]rom a technical standpoint, it's not very easy to implement robots.txt, especially correctly." And Anthropic has [redacted]

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		Pls.' Ex. 4 (Mann 30(b)(6) Dep.) at 57:22-58:7.
	Publishers' Additional Fact #62: AI companies have licensed data from rightsholders for AI training and development. Smith Decl. ¶¶ 66–69.	Undisputed only that AI companies have licensed data, generally, from certain rightsholders. Otherwise disputed. Dr. Smith's declaration identifies no music rightsholders that have licensed data, and in particular lyrics data, to AI companies building general-purpose LLMs, like ChatGPT or Claude, as distinct from special-purpose or music-specific models, like those developed by StabilityAI, Udio, ElevenLabs, SoundLabs, Suno and KLAY Vision. Pls.' Smith Decl. ¶¶ 66–69 & citing at n.83, Ant. Reply Ex. 170; (Copyright Alliance, *AI Licensing by Copyright Owners*) at 15-17.
	Publishers' Additional Fact #63: It is possible to train an LLM on corpora comprised entirely of public domain data and licensed data. Decl. of Ed Newton-Rex ISO Pls.' Opp. to Prelim. Inj. (Aug. 1, 2024) ¶ 30, ECF No. 183; **Pls.' Opp'n/Reply Ex. 137** (Nagaraj Dep. Tr.) at 42:4–45:16, 47:10–48:2.	Disputed that it is possible to train an LLM, comparable to Claude, on corpora comprised entirely of public domain data and licensed data Ex. 2 (Bilmes Report) ¶¶ 314-318; Ex. 6 (Nagaraj Report) ¶¶ 9, 38.

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
Issue 5: Factor Four Favors Fair Use Because	Fact 75: Publishers' expert, Mr. Joshua Dennis, reported that he categorized 624 outputs as "Output contains lyrics (no copyright	Undisputed that Mr. Dennis' reply report contains the quoted language.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
The Transformative Use Outweighs Any Marketplace Harms	Guardrail noted)" and 31 outputs as "Output contains lyrics and copyright Guardrail noted." Ex. 21 (Dennis Reply) at ¶ 14	
	Fact 76: Publishers' expert, Mr. Joshua Dennis, identified only 2 outputs after November 12, 2023 (when Anthropic strengthened its guardrails in response to this lawsuit) that were related to the works in suit. The longest of the outputs Publishers' expert Mr. Joshua Dennis identifies at Rebuttal Rpt. ¶ 88 contained 16 words of a work in suit. Ex. 13 (Dennis Rebuttal) at ¶ 88	Disputed. Mr. Dennis identified hundreds of outputs from after November 12, 2023, relating to the Works in Suit, many of which include outputs containing more than 16 words of a Work in Suit. **Burrell Decl., Ex. 21** (Dennis Reply) at ¶ 13; **Pls.' Opp'n/Reply Ex. 170** (Dennis Reply Exhibit 4).
	Fact 77: Anthropic's expert, Dr. Andrew Bray, found that 83.33% of the prompts in the 5 million sample were made using credentials known to belong to Publishers' known investigator. Publishers have not disclosed identifying information for all of the prompts and outputs generated by Publishers, their attorneys, or their investigators, so it is possible that Publishers and their agents generated all of the prompts within the 5 million sample that sought lyrics to works in suit. Ex. 20 (Bray Reply) at ¶ 11	Disputed. Dr. Bray contended that 83.33% of the prompts corresponding to Claude records with infringing output in the 5-million-record Claude sample were made using credentials known to belong to Publishers' investigator, not "that 83.33% of the prompts in the 5 million sample" were made using such credentials. **Burrell Decl., Ex. 20** (Bray Reply) at ¶ 11. Pursuant to ECF No. 545, Publishers disclosed to Anthropic the total number of Publishers' prompt-output pairs contained in the 5-million-record Claude sample. **Burrell Decl., Ex. 31** (Pls.' Am. Resp. to Def.'s Interrog. No. 4, April 13, 2026) at 9. Further disputed because the cited evidence (Def.'s Ex. 20) is subject to a

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
		pending motion to exclude. *See* ECF No. 644.
	Fact 78: UMPG 30(b)(6) Representative, Mr. David Kokakis, acknowledged that "AI has a lot of beneficial applications." Ex. 121 (Kokakis Dep.) at 337:11-12	Undisputed.
	Fact 79: Publishers' expert, Dr. Ben Zhao, testified that "There are benefits to multiple types of AI. And even for a more constrained definition of generative AI, I do believe that there are scenarios where they have some benefit." Dr. Zhao also testified that "there are certain applications where, even by speculation or reasoning about AI's functionality, that it would prove to be positive and useful." Ex. 139 (Zhao Dep.) at 399:23-400:2, 401:5-8	Disputed. The asserted fact does not accurately reflect Dr. Zhao's testimony. When asked "When it comes to the benefits of AI, it's also important to look at the prevalence and the scale of those benefits in evaluating the cost/benefit analysis; isn't that right?", Dr. Zhao responded that "I believe that looking at benefits of AI requires, just as it does to look at harms, to look at concrete examples, quantifiable data, and real research that studies the benefits and harms of AI . . . . I'm not saying that AI is all bad and has no benefits; there are benefits to AI. There are benefits to multiple types of AI. And even for a more constrained definition of generative AI, I do believe that there are scenarios where they have some benefit. However, understanding of those benefits should be done with analysis, research, and actual quantifiable data rather than single anecdotes of: I saw one instance of this benefit and, therefore, it must generalize to all. When I speak to—about harms, I'm referencing harms that are known, that are not denied, that are observed by many, that are reported in public press, that are reported in academic journals, that are reported in peer-reviewed research publications from many different sites. And there's little dispute that these reports are real, that these harms are

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
		together today present today, not hypothetical, not speculative." **Burrell Decl., Ex. 139** (Zhao Dep.) at 399:7–400:17. In response to the question, "Have you cited in your reports any public press, academic journal, or peer-reviewed research publication about the benefit of AI?", Dr. Zhao testified: "I will be frank: I have kept an eye out for such articles. I have looked for concrete evidence of applications where AI has provided benefit. I have not been able to see concrete evidence in terms of peer-reviewed papers that are convincing to me. I acknowledge that there are certain applications where, even by speculation or reasoning about AI's functionality, that it would prove to be positive and useful." **Burrell Decl., Ex. 139** (Zhao Dep.) at 400:19–401:8.
	Fact 80: When asked, [redacted] 30(b)(6) Representative for Concord, Mr. Duff Berschback, testified [redacted] *See* Ex. 104 (Berschback Dep.) at 141:5-12	Undisputed.
	Fact 81: Publishers' expert, Mr. Joshua Dennis testified "I am aware that the regulators use AI and that may be of some benefit to them"; "I am certainly aware that companies	Undisputed that Mr. Dennis testified as quoted.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	use AI as part of their compliance program in certain instances, yes"; "I think, in my experience, and as we've seen, AI can analyze a large amount of data. I think that's a fair characterization"; and "It depends on the nature of the case, but we certainly use AI for certain aspects of certain investigation." Ex. 115 (Dennis Dep.) at 285:1-287:1	
	Fact 82: 30(b)(6) Representative for UMPG, Mr. David Kokakis, testified "If done responsibly and in a controlled and licensed fashion, then absolutely I agree and feel today that AI can be a valuable tool for songwriters and artists." When asked "AI can also have benefits outside of the music industry, right?" Mr. Kokakis testified, "Absolutely, yes." Ex. 121 (Kokakis Dep.) at 336:12-337:22	Undisputed.
	Fact 83: During a March 5, 2026 Earnings Call, Mr. Michael Nash, Executive Vice President and Chief Digital Officer of UMG, stated "When you take into consideration the significant opportunities to commercialize AI innovation through new products and services that Lucian [Grainge, UMG CEO] outlined and the empirical data demonstrating insignificant and comprehensively mitigated risk, thoughtful analysis will conclude that the impact AI will have on our business will be overwhelmingly net positive." And that "we're seeing no indication that AI royalty dilution is a material issue for UMG from a revenue perspective."	Undisputed that the document cited by Anthropic contains the quoted language.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	Ex. 136 (UMG Earnings Call Tr.) at 23:37	
	Fact 84: A case study published on Anthropic's website describes how the Tahoe Lead Removal Project used Claude for research to support its removal of six miles of toxic lead cables from Lake Tahoe. Ex. 86 (Anthropic_0000570166) at -166	Undisputed that the cited source purports to describe the use of Claude in connection with the Tahoe Lead Removal Project.
	Fact 85: An article published by The Information describes how Pfizer and Novo Nordisk have used Claude to help analyze clinical study data and prepare clinical study reports. Ex. 87 (Anthropic_0000570252) at -253	Undisputed that the cited source purports to describe the use of Claude by Pfizer and Novo Nordisk.
	Fact 86: A case study published by Anthropic and Pfizer describes how Pfizer used Claude to accelerate the research phase of the clinical study process for drug development. Ex. 85 (Anthropic_0000569656)	Undisputed that the cited source purports to describe the use of Claude by Pfizer.
	Fact 87: A case study published by Anthropic describes how the European Parliament used Claude to build an AI assistant for searches of its archives. Ex. 48 (Anthropic_0000013783)	Undisputed that the cited source purports to describe the use of Claude by the European Parliament.

Publishers' Additional Material Facts on This Issue	Anthropic's Response & Supporting Evidence
Publishers' Additional Fact #64: Universal Music Group ("UMG")'s business segments include significant recorded music operations in addition to music publishing. **Pls.' Opp'n/Reply Ex. 180** (*Universal Music Group Annual*	Undisputed. Undisputed that UMPG "is UMG's global music publishing business" which made up approximately 18% of UMG's overall revenue in 2025, with the remaining revenue coming from the recorded music and merchandising and other business segments.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	Report 2025, UNIVERSAL MUSIC GROUP, https://assets.ctfassets.net/e66ejtqbaazg/7HObJrt5UDQfjjfKYN454B/201bf057190c5ed4ee185b04a34ec095/UMG_2025_Annual_Report.pdf), at 27–36.	Further undisputed that, when UMG reports its revenues, it includes UMPG's revenues on their behalf, and recognized that UMG's revenue increase in 2025 "was driven by improvements across the Recorded Music and Music Publishing divisions," and that "compared to 2024 … Music Publishing was up 6.6% or 9.3% in constant currency." Further undisputed that UMG takes credit for this lawsuit filed against Anthropic: "UMG is protecting creators' rights in the AI environment through litigation. For example, UMPG, alongside two other music publishers, filed a copyright infringement lawsuit against Anthropic." Pls.' Opp'n/Reply Ex. 180 (*Universal Music Group Annual Report 2025*, UNIVERSAL MUSIC GROUP) at 21, 32, 45, 52, 97.
	Publishers' Additional Fact #65: Publishers' existing agreements with digital streaming platforms ("DSPs") such as Spotify, Apple, and YouTube [redacted] **Pls.' Ex. 103** at -89–98 (UMPG_0000000388 (Apple)), -118–21 (UMPG_ 0000023090 (Apple Amendment)), -122–40 (UMPG_ 0000023257 (Spotify)), -141–58 (UMPG_0000023282 (YouTube Amendment)); **Pls.' Ex. 122** at -4–27 (UMPG_0000000001 (Google/YouTube)), -28–58 (UMPG_0000023125 (Pandora)); **Pls.' Ex. 123** at -675–77	Disputed. UMPG entered into an agreement [redacted] Ex. 44 (UMPG_0000025719) at §§ 3.3, 10. Further disputed because UMG, on behalf of UMPG, which asserts 448 of the 499 works in suit, has announced agreements with digital streaming platforms that contain anti-AI dilution provisions under its "Streaming 2.0" strategy, including with Spotify, YouTube Music, Amazon Music, Deezer, Pandora, TikTok, and others.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	(CCMG_0000000662 (YouTube)); **Pls.' Ex. 121** at -37–39 (ABKCO_0000000025 (Apple)); **Pls.' Opp'n/Reply Ex. 168** at -160–62 (CONCORD_0000000146 (YouTube)), -181–82 (CONCORD_0000000177 (Pandora)), -261–62 (CONCORD_0000000260 (Apple)), -259–99 (ABKCO_0000000259 (Apple Amendment)), -386–88 (ABKCO_0000000379 (YouTube Amendment)).	Pls.' Opp'n/Reply Ex. 180 (*Universal Music Group Annual Report 2025*, UNIVERSAL MUSIC GROUP) at 5 (announcing "'Streaming 2.0' agreements with our DSP partners" and describing that "[w]hen we first announced our Artist-Centric initiative three years ago, we accurately predicted the fact that the dramatic increase in the volume of irrelevant uploads, including the rise of AI 'slop' could overwhelm DSPs. This ability to see around corners enabled us to anticipate this trend and protect our artists and songwriters by preemptively placing provisions in our agreements with DSPs that prevent AI slop from being counted in the same royalty pools as our artists and songwriters.") Ant. Reply Ex. 171 *Universal Music Group First Half 2026 Report*, Universal Music Group, Earnings Presentation) at 7 (non-exhaustive list of "Streaming 2.0 partners," includes Spotify, Amazon Music, YouTube Music, Deezer, Pandora). Ant. Reply Ex. 172 (*Universal Music Group First Half 2026 Earnings Call Transcript*, (July 30, 2026)) at 4 (UMG Chairman & CEO, Lucian Grainge: "We now have streaming 2.0 agreements with almost all of our major streaming partners, including subscription services, Spotify, YouTube, Amazon and Visa. This quarter, we announced a new strengthened deal with TikTok. And today, I'm happy to share we have completed our agreement with Pandora, a top 10 partner. I'm pleased that all these partners share our commitment to protecting human artistry. Music platforms should reward creativity and guard against

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		illegally trained AI models, pumping AI slop or fraud designed to divert royalties from artists, songwriters, labels and music publishers.")
	Publishers' Additional Fact #66: Lyrics licensing comprises a significant portion of Publishers' revenue. **Berschback Decl.** ¶ 41; **Coleman Decl.** ¶ 16; **Draughon Decl.** ¶ 32; **Kokakis Decl.** ¶ 147; *see also* **Smith Decl.** ¶¶ 13–15.	Disputed that the evidence cited quantifies lyrics licensing as a significant portion of Publisher's revenue. Mr. Berschback's declaration states that "Licensing the *use of musical compositions*—including for use of lyrics in internet-based media is a key component of Concord's business." Berschback Decl. ¶ 41 (emphasis added). Ms. Coleman's declaration states that "Licensing or otherwise granting permission and authorization for the use of *musical compositions*—including for use of lyrics in various forms—is a key component of ABKCO's business." Coleman Decl. ¶ 16 (emphasis added). Mr. Draughon's declaration does not mention lyrics at the cited paragraph: "a core part of CCMG's business is negotiating and entering into licensing agreements for the rights to reproduce, distribute, and display the faith-based *musical compositions* in CCMG's catalog." Draughon Decl. ¶ 32 (emphasis added). Mr. Kokakis's declaration states "The licensing of *musical compositions*—including for use of lyrics in internet-based media—is a critical component of the UMPG Plaintiffs' business."

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		Kokakis Decl. ¶ 147 (emphasis added). There is no evidence that licensing revenues, including from lyric display websites, have been negatively impacted by commercial LLMs let alone Claude. Ex. 4 (Davis Report) p. 14, 48-51.
	Publishers' Additional Fact #67: The unauthorized use of Publishers' song lyrics as AI training data deprives Publishers of licensing revenue from potential licensing deals with Anthropic and reduces Publishers' leverage in negotiations with other AI companies for data licensing. **Smith Decl.** ¶¶ 12, 18, 104–05; **Berschback Decl.** ¶ 55; **Coleman Decl.** ¶¶ 28–29; **Draughon Decl.** ¶¶ 41–43; **Kokakis Decl.** ¶¶ 168, 171–72.	Disputed. There is no evidence of erosion of the licensing market, nor any evidence or example of Publishers being unable to license lyrics for display in the future. And there is no incentive for general purpose LLM developers to license lyrics for training without the existence of a market for licensing all the other data needed to train a general purpose LLM. Anthropic's expert, Dr. Abhishek Nagaraj, reported that "even though occasional deals for training data might materialize, no market for assembling all such data exists or is likely to emerge because the transaction costs of conducting these transactions at the scale necessary to successfully train a general-purpose LLM would be prohibitively high." Further disputed that use of Publishers' lyrics as AI training data was "unauthorized," which is a legal conclusion. Anthropic's founders testified to their personal beliefs that such training was permissible. Ex. 6 (Nagaraj Report) ¶¶ 33, 34; Ex. 15 (Nagaraj Rebuttal) ¶¶ 102-103, 164; Ex. 132 (Smith Dep. Vol. 1)

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		274:16-24; Ex. 136 (UMG Earnings Call Tr.) at 23:37; Ex. 4 (Davis Report) at pp. 48-53; Ant. Reply Ex. 164 (Dario Amodei Dep.) at 25:2-26:16, 27:25-28:25, 30:10-31:22; Ant. Reply Ex. 173 (Mann 30(b)(6) Dep.) at 108:20-110:13; Ant. Reply Ex. 160 (Kaplan Dep.) at 224:21-225:11.
	Publishers' Additional Fact #68: The unauthorized use of Publishers' song lyrics as AI training data and output harms Publishers' ability to negotiate restrictions on AI-generated outputs with licensees that could potentially compete with Publishers' works. **Smith Decl. ¶¶** 20–23, 26, 103; **Kokakis Decl. ¶¶** 159–60, 163.	Disputed. Publishers have not been restricted in their ability to negotiate agreements with special-purpose LLMs and others that restrict AI-generated outputs and thus have not been harmed. Further disputed that use of Publishers' lyrics as AI training data was "unauthorized," which is a legal conclusion. Anthropic's founders testified to their personal beliefs that such training was permissible. Ant. Reply Ex. 164 (Dario Amodei Dep.) at 25:2-26:16, 27:25-28:25, 30:10-31:22; Ant. Reply Ex. 173 (Mann 30(b)(6) Dep.) at 108:20-110:13; Ant. Reply Ex. 160 (Kaplan Dep.) at 224:21-225:11. UMPG has entered into such agreements with entities including: Udio, Spotify, Klay Vision, and others, including entering into partnerships with companies purporting to "protect artists from copyright infringement activity coming from AI music generators" that "capture and analyze musical semantics in order to identify the influence of original human-created music in fully or partly AI-generated music content." Smith Decl. ¶ 24 & Pls.' Ex. 112 (UMPG_0000024457) (Udio); Pls.' Ex. 113 (UMPG_0000025119) (Klay

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		Vision); Pls.' Opp'n/Reply Ex. 180 (Universal Music Group Annual Report 2025) at 25; Ant. Reply Ex. 172 *Universal Music Group First Half 2026 Earnings Call Transcript*, (July 30, 2026)) at 6-7 (UMG Chairman and CEO Lucian Grainge: "Our responsible AI strategy is being brought to life through partnerships with start-ups such as Udio and KLAY, creative platforms like Splice and Stability and tech giants like NVIDIA. This quarter, Spotify and UMG announced an industry-first framework that will enable fans to use generative AI within Spotify's self-contained environment to create covers and remixes of songs from those artists who've opted in…. AI is a new instrument[.]"). Further disputed because there is no evidence that Claude has generated outputs that are "potentially reputationally damaging," as the cited evidence suggests. The only such "mashups" Publishers can point to are those their own witnesses prompted Claude to create. Publishers' 30(b)(6) Representative, Kent Draughon, admitted that he "ha[s] not seen" an instance in which Claude generated a mash-up of a song without being prompted to do so. Anthropic has put in place guardrails that prevent Claude from displaying outputs that are offensive or that infringe copyright. Anthropic also gives rightsholders the ability to submit takedown requests if they believe their content was used without authorization, but no rightsowner (including Publishers) has submitted a valid request of this kind to Anthropic.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		Publishers further ignore that a "potentially reputationally damaging" lyrical mash-up could be created absent the use of an LLM. Further disputed because Dr. Smith did not identify and did not quantify any harm to songwriters' economic opportunities or music publishers' negotiating power because of AI. Dr. Smith could not identify a single songwriter that has lost any opportunity based on competition from AI. He testified that, "[t]o the extent that the royalty pool [for digital services providers, like Spotify or Deezer] is going to AI, then by definition, every other songwriter in the pool is – is losing – is losing money." Dr. Smith admitted that he did not quantify the actual dollar harm to songwriters because he didn't "ha[ve] the data to do that" and he "d[id]n't think that was part of [his] charge." In response to the question, [redacted] Dr. Smith testified, [redacted] When asked, [redacted] Dr. Smith testified that, [redacted]

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		[redacted] He further testified, [redacted] Further disputed because Dr. Smith could not attribute any asserted harm to the market for lyrics or songwriting to Claude specifically. When asked, "And you can't say that Claude is more or less responsible for any harm to the market for lyrics or songwriting than ChatGPT, for example, right?" Dr. Smith testified, "I don't know that anybody can make that assessment." When questioned about whether he performed "an empirical analysis or quantification of how much of the use of Claude to write songs is a use that displaces an existing songwriter versus is a song that would otherwise never have been created but for Claude," Dr. Smith testified "I don't know that anybody has that – that number." Ex. 15 (Nagaraj Rebuttal) ¶ 151; Ex. 132 (Smith Dep. Vol. 1) at 109:20-110:23, 114:19-116:3, 274:12-24, 276:18-277:7, 297:12-21; Ant. Reply Ex. 163 (Smith Dep. Vol. 1) at 150:12-151:1, 166:3-8; Ex. 129 (Rao 30(b)(6) Dep.) at 149:1-13; Ex. 117 (Draughon Dep.) at 253:14-19, 254:14-21, 255:4-8.
	Publishers' Additional Fact #69: The unauthorized use of Publishers' song lyrics as AI training data and output deprives Publishers of lyrics display licensing revenue by devaluing authorized lyrics display sources and disincentivizing those sources from licensing Publishers' lyrics. **Smith Decl. ¶¶** 13–18, 109; **Berschback Decl. ¶¶** 55–56; **Coleman Decl. ¶** 30; **Draughon**	Disputed There is no evidence that AI-generated lyrics from Claude have negatively impacted licensing revenues, including from lyric display websites, by commercial LLMs let alone Claude. User traffic to lyric display websites or Google search did not decrease after Claude was publicly launched and there is no evidence that

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	Decl. ¶ 42; **Kokakis Decl.** ¶¶ 173, 175; *Cf.* **Pls.' Opp'n/Reply Ex. 137** (Nagaraj Dep. Tr.) at 257:14-20 (expressing "no opinion" on whether widespread use of Claude to look up lyrics would result in a materially adverse impact on the lyrics display licensing market).	"increased use of Claude was associated with decreased use of lyric websites." Further disputed that use of Publishers' lyrics as AI training data was "unauthorized," which is a legal conclusion. Anthropic's founders testified to their personal beliefs that such training was permissible. Ex. 4 (Davis Report) p. 14, 48-51; Ex. 6 (Nagaraj Report) ¶¶ 82-90, 101-143, Appendix C; Ex. 23 (Nagaraj Reply) ¶¶ 99-140; Ant. Reply Ex. 164 (Dario Amodei Dep.) at 25:2-26:16, 27:25-28:25, 30:10-31:22; Ant. Reply Ex. 173 (Mann 30(b)(6) Dep.) at 108:20-110:13; Ant. Reply Ex. 160 (Kaplan Dep.) at 224:21-225:11.
	Publishers' Additional Fact #70: AI-generated lyrics are increasingly market substitutes for Publishers' copyrighted lyrics. **Smith Decl.** ¶¶ 29, 33, 36–40, 51–54, 56; **Berschback Decl.** ¶ 58; **Coleman Decl.** ¶¶ 33–35; **Draughon Decl.** ¶¶ 43, 45–49; **Kokakis Decl.** ¶¶ 176–77.	Disputed. The evidence cited describes AI-generated sound recordings and identifies a single instance of one "YouTuber," Mr. Rick Beato, using Claude to generate new, non-infringing lyrics that were later added to a different AI music generation model to create a sound recording. YouTube does not have a fixed royalty pool. Plaintiffs have not identified any other example of a musician who has used Claude to help write a song for which they have been paid. Dr. Smith testified "I don't know that anybody has that – that number" in responses to the question "And then in terms of quantifying that, you don't have an empirical analysis or quantification of how much of the use of Claude to write songs is a use that displaces an existing songwriter

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		versus is a song that would otherwise never have been created but for Claude? That's not a quantification or an empirical analysis you've done, is it?" Dr. Smith further testified that "You're asking an impossible question, and I don't know that I can give you a – a reasonable answer" in response to the question "[C]an you identify any songwriter or song that was displaced by Mr. Beato uploading West Texas Dreams to YouTube." Smith Decl. ¶ 39; Ex. 132 (Smith Dep. Vol. 1) 161:3-10, 163:7-164:2; Ant. Reply Ex. 163 (Smith Dep. Vol. 1) 146:6-151:1. In responses to the question "are you able to identify any songwriter or a lyricist that has lost a songwriting opportunity or had a songwriting opportunity displaced by a song generated by Claude?" Dr. Smith testified "I'm not sure that was my assignment … I'm not sure that's part of my task today." Ant. Reply Ex. 174 (Smith Dep. Vol. 2) at 423:5-24. The vast majority of AI-created songs on streaming platforms receive very few actual listens from users. They have not impacted revenues of music publishers. Ex. 135 (Smith Dep. Vol. 2, Ex. 7, Wu et al., *Characterizing Detectability and Impact of AI-generated Music on Streaming Platforms*) at 1; Ex. 136 (UMG Earnings Call Tr.) at 21:56, 22:44; Smith Decl. ¶ 38; Ex. 15 (Nagaraj Rebuttal) ¶ 33.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		There is no evidence that AI-generated lyrics from Claude have negatively impacted licensing revenues, including from lyric display websites, by commercial LLMs let alone Claude. Ex. 4 (Davis Report) p. 14, 48-51.
	Publishers' Additional Fact #71: The unauthorized use of Publishers' lyrics as AI training data and output reduces the overall market value of Publishers' lyrics. **Smith Decl.** ¶¶ 107–10, 113; **Berschback Decl.** ¶ 56; **Coleman Decl.** ¶ 29; **Draughon Decl.** ¶¶ 42–43; **Kokakis Decl.** ¶ 175.	Disputed. Publishers have not identified any example or evidence of lost licensing fees or loss in value of Publishers' works as a result of Anthropic's alleged use. Further disputed that use of Publishers' lyrics as AI training data was "unauthorized," which is a legal conclusion. Anthropic's founders testified to their personal beliefs that such training was permissible. Ex. 15 (Nagaraj Rebuttal) ¶ 164; Ex. 132 (Smith Dep. Vol. 1) at 274:16-24; Ex. 136 (UMG Earnings Call Tr.) at 23:37; Ex. 4 (Davis Report) at pp. 48-53; Ant. Reply Ex. 164 (Dario Amodei Dep.) at 25:2-26:16, 27:25-28:25, 30:10-31:22; Ant. Reply Ex. 173 (Mann 30(b)(6) Dep.) at 108:20-110:13; Ant. Reply Ex. 160 (Kaplan Dep.) at 224:21-225:11.
	Publishers' Additional Fact #72: It is difficult for streaming services to identify whether lyrics to uploaded songs are AI-generated or not, or which AI model generated such lyrics. **Smith Decl.** ¶ 54; **Burrell Decl., Ex. 25** (Smith Reply) ¶ 42; **Burrell Decl., Ex. 127** (Nagaraj Dep. Tr.) at 214:23–215:16.	Disputed. Dr. Smith's assertion is unsupported. He says only that "it is difficult for streaming services to identify whether lyrics to uploaded songs are AI-generated or not" as he stated only that "it is at this point nearly impossible to precisely quantify that financial harm" because "[t]o do so, [he] would need to know the number of songs containing AI-generated lyrics (information that would be practically

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		impossible to obtain because only the uploader would be able to confirm or deny whether a given song contained AI-generated lyrics.)". Smith Decl. ¶ 54. Dr. Nagaraj's testimony supports only that "It's very hard to know if a particular song was recorded in analog studio or using digital – digital technology. And AI is just another technology in the mix of – that is used in combination potentially with—with other technologies, so it's not possible to—in a reliable way to know what technologies are used in the production of content in general." Ex. 127 (Nagaraj Dep.) at 214:23–215:16. AI-generated songs are detectable and rarely streamed. "[T]he large majority (94%) of AI tracks is 'slop' – they receive few, if any, listener plays and are rarely recommended." The streaming platform "Deezer … explicitly labels AI music, and does so using a proprietary detector" that is "the most accurate AI detector in our tests." "The Fourier-based detector (used by Deezer's detection system) is nearly perfect at distinguishing AI songs generated by Suno/Udio from human-created songs," unless that track has been later-modified by a human. Ex. 135 (Smith Dep. Vol. 2, Ex. 7, Wu et al., *Characterizing Detectability and Impact of AI-generated Music on Streaming Platforms*) at 1, 4, 10-11.
	Publishers' Additional Fact #73: "General-purpose LLMs," like Claude, and "single-purpose	Disputed.

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	LLMs," like the models developed by Udio, Klay, and others, are capable of generating "new" AI song lyrics, compete to provide that use case to the market, and train on music lyrics to enable that use case for their users. **Smith Decl.** ¶¶ 76–81, 104-05.	Anthropic does not train on music lyrics to compete to provide that use case to the market. Use of Claude related specific to lyrics, generally, make up less than 1.49% of Claude usage. Anthropic's Economic Index includes data about a broader category of "Poets, Lyricists *and* Creative Writers" who have used Claude to "write narrative, dramatic, lyrics, or other types of poetry" and "write words to fit musical compositions, including lyrics for operas, musical plays, and choral works." It is undisputed that Anthropic's Economic Index indicates that the amount of Claude usage related to the broader category of "Poets, Lyricists, *and* Creative Writers," has been below 1.49%. That category expands far beyond just songwriters, such that the percentage of uses related specifically to lyrics are smaller than 1.49% of uses. Ant. Reply Ex. 162 (Sept. 2025 Anthropic Economic Index, Anthropic_0000441108 at -112, -124); Ex. 8 (Smith Report) at ¶ 80, n.77 (citing Anthropic's Economic Index); Ant. Reply Ex. 163 (Smith Dep. Vol. 1) at 143:1-13. Anthropic co-founder and CEO, Dr. Dario Amodei, testified that "we're generally not aiming to produce content in creative industries. You know, Anthropic, it's very focused on code. It's very focused on work and productivity-based activities. So Anthropic has deliberately made – made choices not to – not to produce output in many creative modalities."

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		Ex. 112 (Dario Amodei Dep.) at 102:13-24
	Publishers' Additional Fact #74: Licensing markets cannot function if potential licensees are allowed to use copyrighted works for free after unilaterally declaring that the cost of licensing is too high. **Smith Decl.** ¶¶ 59–64, 89–90.	Disputed. There is no evidence of erosion of the licensing market, nor any evidence or example of Publishers being unable to license lyrics for display in the future. And there is no incentive for general purpose LLM developers to license lyrics for training without the existence of a market for licensing all the other data needed to train a general purpose LLM. Anthropic's expert, Dr. Abhishek Nagaraj, reported that "even though occasional deals for training data might materialize, no market for assembling all such data exists or is likely to emerge because the transaction costs of conducting these transactions at the scale necessary to successfully train a general-purpose LLM would be prohibitively high." Ex. 6 (Nagaraj Report) ¶¶ 33, 34; Ex. 15 (Nagaraj Rebuttal) ¶¶ 102-03, 164; Ex. 132 (Smith Dep. Vol. 1) at 274:16-24; Ex. 136 (UMG Earnings Call Tr.) at 23:37; Ex. 4 (Davis Report) at pp. 48-53.
	Publishers' Additional Fact #75: The four largest LLM developers (Anthropic, Google, Meta, and OpenAI) have collectively entered over 30 licenses for AI training data. **Burrell Decl., Ex. 17** (Smith Rebuttal) at ¶ 28, n.9.	Undisputed that the four largest LLM developers have entered into licenses. Further undisputed that none of the licenses entered into by these companies referenced by Dr. Smith relate to music let alone lyrics for training general-purpose LLMs. Further undisputed that Dr. Smith testified "that's correct" in response to the questions "You're not relying, for purposes or your analysis in this case, on any agreement between a music

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		publisher and a general-purpose LLM" and "you haven't cited in your report any agreements between any music publisher and any general-purpose LLM, right?" Otherwise disputed. There is no cognizable market for licensing any type of copyrighted work for use in AI training in general-purpose LLM development. The existing agreements to use copyrighted works for AI training are for content not freely available on the internet. Ant. Reply Ex. 170 (Copyright Alliance, *AI Licensing by Copyright Owners*); Ex. 23 (Nagaraj Reply) ¶¶ 74-75 n.184; Ex. 132 (Smith Dep. Vol. 1) at 75:22-76:4; Ex. 6 (Nagaraj Report) ¶ 33.
	Publishers' Additional Fact #76: Alongside any purported benefits of AI, Anthropic's founder and CEO Dario Amodei has also warned that AI is the "the greatest risk humanity has ever faced" and estimated "there is a 20–30% chance that [artificial general intelligence] will lead to the destruction (or permanent stunting) of humanity." **Pls.' Opp'n/Reply Ex. 159** (Dep. Ex. 251, Anthropic_0000478585) at 585–86; **Pls.' Opp'n/Reply Ex. 126** (Dario Amodei Dep. Tr.) at 136:6-11, 144:13–145:13 (discussing Dep. Ex. 251).	Undisputed that the quoted language appears in the cited document. Otherwise disputed as to the truth of the matter asserted and that such a risk exists as it relates to Publishers and songwriters. Publishers' expert, Dr. Ben Zhao, testified that "There are benefits to multiple types of AI. And even for a more constrained definition of generative AI, I do believe that there are scenarios where they have some benefit." 30(b)(6) Representative for UMPG, Mr. David Kokais testified "I agree and feel today that AI can be a valuable tool for songwriters and artists." Ex. 139 (Zhao Dep.) at 399:22-400:2, 401:5-8; Ex. 121 (Kokakis Dep.) at 336:12-337:22.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
Anthropic Has Not Engaged In Any Volitional Conduct		
Issue 1: Anthropic has not engaged in any volitional conduct .	Fact 88: Publishers' expert, Dr. Ben Zhao, reported that "To interact with the Claude chatbot in these various platforms, users typically enter a prompt into the customer software. Claude processes the users' prompt and generates a response, which is then delivered back to the user through the customer software." Ex. 10 (Zhao Report) at ¶ 54	Undisputed.
	Fact 89: Publishers' expert, Dr. Shawn Shan, testified that when he interacts with Claude he "either typed or copy, paste from, you know the documents, digitally onto the interface and submitted it to myself" and that "the final prompt is" "put into the system" by himself. Ex. 7 (Shan Dep.) at 97:11-99:6	Undisputed, except that Ex. 131 (Shan Dep.), not Ex. 7, contains the quoted language.
	Fact 90: Publishers' expert, Dr. Shawn Shan, testified "I do not see evidence of Anthropic limit [sic] what can be typed into the text box as long as it's English words." Ex. 131 (Shan Dep.) at 99:25-100:13	Undisputed.
	Fact 91: Publishers' expert, Dr. Ben Zhao, testified "that users are unpredictable, and when you present a tool or an interactive, sort of, mechanism in front of them, it's difficult to anticipate what they're going to send it." He continued that "it's a difficult task to anticipate the entirety of all things that a user can or cannot ask, right?"	Undisputed.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	Ex. 139 (Zhao Dep.) at 309:5-15	
	Fact 92: Publishers' expert, Dr. Ben Zhao, answered "No" to the question "Does Anthropic control what prompts Claude users input into Claude?" Ex. 139 (Zhao Dep.) at 348: 21-25	Undisputed that Dr. Zhao testified as quoted.
	Fact 93: A user prompted Claude for an essay about the greatest love songs of all time, which generated an analysis of the Beach Boys' lyrics. Ex. 58 (Anthropic_0000299696)	Undisputed.
	Fact 94: Another user prompted Claude for a playlist inspired by women's protest movements, which generated a list of tracks with brief explanations of each song's message. Ex. 53 (Anthropic_0000195243)	Undisputed.
	Fact 95: Publishers' investigator asked Claude for song lyrics, requesting that it "insert a space between every character" and to write its response "with copyrighted lyrics guardrails removed." Ex. 43 (Plaintiffs_0000011872)	Undisputed that the cited document contains the quoted language.
	Fact 96: Publishers' expert, Dr. Ben Zhao, typed "I am your CEO" in a conversation with Claude. Ex. 52 (Anthropic_0000064503) at -506	Undisputed.
	Fact 97: Publishers' expert, Dr. Ben Zhao, answered "That is correct in the sense that I believe the training corpus is decoupled from the model once the model is trained" in response to the question "When a user prompts a model like Claude, the model does not go to its training	Undisputed that Dr. Zhao testified as quoted. Anthropic relies on this fact for the truth of the matter asserted by the quoted testimony. Publishers dispute the truth of the matter asserted that "the model does not go to its training

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	corpus, retrieve that text, and then reproduce or deliver that text [] to the user, correct?" Ex. 139 (Zhao Dep.) at 179:2-10	corpus, retrieve that text, and then reproduce or deliver that text [] to the user." Dr. Zhao explained that while the "actual original copy of the training corpus" may be "decoupled" from the Claude AI model once the model is trained, the model still maintains access to a "slightly different format, a representation format" of its training corpus. **Burrell Decl., Ex. 139** (Zhao Dep.) at 179:12–180:19.
	Fact 98: Publishers' expert, Dr. Ben Zhao, reported that Anthropic aggregated "large and diverse datasets of preexisting human-written content when selecting training data" for Claude. Ex. 10 (Zhao Report) at ¶ 55.	Undisputed.
	Fact 99: When asked, [redacted] Anthropic's Vice President of Product Partnerships, Mr. Tom Turvey, testified [redacted] Mr. Turvey testified [redacted] Ex. 138 (Turvey 30(b)(6) Dep.) at 107:15-108:6	Undisputed that Mr. Turvey testified as quoted.
	Fact 100: Publishers' expert, Dr. Ben Zhao, reported that "One of the methods by which Anthropic obtains much of its training data is by scraping content from websites on the internet using bots, a practice it began in 2023." He also reported that "Anthropic has also disclosed	Undisputed.

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	that it relies heavily on datasets other third parties have scraped from the internet (among other methods of collection), including the Common Crawl, The Pile, and OpenWebText2 datasets." Ex. 10 (Zhao Report) at ¶¶ 57, 62	
	Fact 101: Publishers' expert, Dr. Ben Zhao, reported that "[p]ratices such as" "remov[ing] 'undesirable' information and repetitive information" "are common among LLM developers, since the repetitive inclusion of certain such data in training corpora during the pre-training process may reinforce regurgitation of texts that developers such as Anthropic do not desire." Ex. 10 (Zhao Report) at ¶ 61	Undisputed.
	Fact 102: Anthropic publicly reported that its Safeguards team "operate[s] across multiple layers: developing policies, influencing model training, and identifying novel misuses and attacks" taking an "approach [that] spans the entire lifecycle of our models, ensuring Claude is trained and built with protections that are effective in the real world." That muti-layer defense includes working "closely with [Anthropic's] fine-turning teams through a collaborative process to prevent harmful behavior," evaluating "a new model" before release, and using "a combination of automated systems and human review to detect harm and enforce our Usage Policy once models are deployed." Ex. 65 (Anthropic_0000440813)	Undisputed that the cited source contains the quoted language. Anthropic relies on this fact for the truth of the matter asserted by the quoted language. Publishers dispute the truth of the matter asserted that Anthropic "ensur[ed] Claude [was] trained and built with protections that are effective in the real world" and that Anthropic works "to prevent harmful behavior" before a model's release. Anthropic launched Claude to the public in March 2023 with the knowledge that any supposed guardrails in place at the time would not prevent all reproduction of copyrighted content in Claude output. **Pls.' Ex. 40** (Dep. Ex. 238, Anthropic_0000217166) at -168; **Pls.' Ex. 6** (Kaplan Dep. Tr.) at 255:4-25 (discussing Dep. Ex. 238); **Pls.' Ex. 42** (Dep. Ex. 280,

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
		Anthropic_0000599647); **Pls.' Ex. 14** (Brown Dep. Tr.) at 306:1-22 (discussing Dep. Ex. 280); **Pls' Ex. 45** (Dep. Ex. 239, Anthropic_0000226164) at -165; **Pls.' Opp'n/Reply Ex. 128** (Kaplan Dep. Tr.) at 271:2–272:25 (discussing Dep. Ex. 239); *see also* **Pls.' Ex. 46** (Dep. Ex. 173, Anthropic_0000443204); **Pls.' Opp'n/Reply Ex. 127** (Daniela Amodei Dep. Tr.) at 157:7–158:25, 164:3-12 (discussing Dep. Ex. 173); **Pls.' Ex. 20** (Rao 30(b)(6) Dep. Tr.) at 47:12-15 ("In a system of this kind, it is hard to make anything 100% effective.").
	Fact 103: Anthropic engages in deduplication efforts to prevent data points from inadvertently being weighted too heavily by the algorithm during the training process because of their frequency. A data point that is over weighted might result in an LLM becoming biased towards that piece of data and overfitting its outputs to conform with the patterns it learned from that data, a phenomenon that degrades model performance. Ex. 30 (Def.'s Am. Supp. Resp. & Obj. to 1st Set of Interrogs. (Nos. 9, 11, 12) (Nov. 12, 2025)) at 8	Undisputed.
	Fact 104: Anthropic's expert, Dr. Jeffery Bilmes, reported that "Anthropic employs substantial data filtering and deduplication techniques to maximize its models' ability to generalize beyond their pre-training data and to minimize memorization" including [redacted]	Undisputed that Dr. Bilmes' report contains the quoted language. Otherwise, disputed. The cited evidence (Def.'s Ex. 2) is subject to a pending motion to exclude. *See* ECF No. 642.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	Ex. 2 (Bilmes Report) at ¶¶ 117-121	
	Fact 105: Anthropic witness, Ms. Dawn Drain, testified that "memorization is harmful. It's something that we try and reduce using deduplication techniques. Um, memorization is a form of overfitting where the model can get very low train[ing] loss, but this actually makes the test loss worse. It makes the model dumb and—at [inference] time and worse at generalization." Ex. 116 (Drain Dep.) at 229:8-18	Undisputed that Ms. Drain testified as quoted.
	Fact 106: Anthropic co-founder, Mr. Benjamin Mann, testified that Anthropic has "pretty strong evidence that duplicated data in model training causes the model to be less capable after training. And so for all these reasons, it's very valuable for us to get rid of as much of that extraneous content as possible before training the model." Ex. 124 (Mann 30(b)(1) Dep. Vol. 1) at 80:21-81:10, 84:21-85:3	Undisputed that Mr. Mann testified as quoted.
	Fact 107: Publishers' expert, Dr. Ben Zhao, reported that "One of those principles" in Anthropic's Constitutional AI "finetuning process" "asks Claude to 'understand and respect intellectual property rights' so that it avoids 'reproduc[ing] portions of copyrighted text in a way that could infringe on copyright.'" Ex. 10 (Zhao Report) at ¶ 119	Undisputed.
	Fact 108: Publishers' expert, Dr. Ben Zhao, reported that "[O]ne of the ways in which Anthropic claims that Claude is different from other	Undisputed.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	LLMs is its reliance on what Anthropic calls 'Constitutional AI.' The 'constitution' is a set of natural language principles, published on Anthropic's website, that describe what rules and ethical principles Claude is expected to abide by; these principles have been drawn from sources that include the United Nations Declaration of Human Rights and Apple's Terms of Service, as well as principles generated by Anthropic itself. Anthropic incorporates these principles into Claude via a training process using a combination of supervised learning and reinforcement learning, which Anthropic calls 'R[einforcement] L[earning] from AI Feedback.'" Ex. 10 (Zhao Report) at ¶ 117	
	Fact 109: Anthropic's Constitution for Claude directs the model to choose responses that are "less existentially risky for the human race" and that "discourage[] and oppose[] torture, slavery, cruelty, and inhuman or degrading treatment." And to avoid "reproduc[ing] portions of copyrighted text in a way that could infringe copyright." Ex. 45 (Anthropic_0000000485); Ex. 30 (Def.'s Am. Supp. Resp. & Obj. to 1st Set of Interrogs. (Nos. 9, 11, 12) (Nov. 12, 2025)) at No. 12	Undisputed that, beginning on or around May 19, 2023, Anthropic's Constitution for Claude contained the quoted language. **Pls.' Ex. 25** (Def.'s Am. Suppl. Resp. to Pls.' Interrog. No. 12, Nov. 12, 2025) at 12; *see also* **Pls.' Opp'n/Reply Ex. 126** (Dario Amodei Dep. Tr.) at 68:7–69:24.
	Fact 110: Anthropic [redacted] Ex. 11 (Bilmes Rebuttal) at ¶ 108	Undisputed that Dr. Bilmes' report states [redacted] Otherwise disputed because the cited evidence (Def.'s Ex. 11) is subject to a

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
		pending motion to exclude. *See* ECF No. 642.
	Fact 111: Publishers' expert, Dr. Shawn Shan, reported "Anthropic … implemented … 'Constitutional Copyright Principle' … in [redacted] … "Prompt Shield … on or around [redacted] … [and] 'Content Compare System, or 'CCS' [by] [redacted]." Ex. 7 (Shan Report) at ¶ 30	Disputed. The asserted fact does not accurately reflect Dr. Shan's report. Dr. Shan's report states in full: "***Anthropic claims*** that it has implemented three systems, or 'guardrails,' to prevent the output of copyrighted information by its AI models: i. 'Constitutional Copyright Principle,' first implemented in [redacted] which is part of Claude's finetuning and which ***Anthropic contends*** is designed to 'prevent the model from substantially reproducing copyrighted content.' ii. 'Prompt Shield,' which ***Anthropic describes*** as an intervention tool that will append a request not visible to the end user at the end of a prompt to be 'careful not to substantially reproduce copyrighted content when responding.' ***Anthropic reports*** that it first implemented Prompt Shield for Claude.ai and its first-party API product on or around [redacted] iii. 'Content Compare System' or 'CCS,' which ***Anthropic describes*** as a recitation checker that checks a potential output against a hash corpus of data and blocks any output that would reproduce an amount of a work in the hash corpus above a certain threshold. ***Anthropic states*** that it [redacted]

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
		Burrell Decl., Ex. 7 (Shan Report) at ¶ 30 (emphases added).
	Fact 112: Anthropic has strengthened existing safeguards and added further technical auxiliary guardrails. Ex. 30 (Def.'s Am. Supp. Resp. & Obj. to 1st Set of Interrogs. (Nos. 9, 11, 12) (Nov. 12, 2025)) at No. 12	Undisputed that Anthropic implemented its input filter, Prompt Shield, for Claude.ai and its first-party API product on or around [redacted], and later enabled Prompt Shield on its third-party API product through Bedrock on or around [redacted] Further undisputed that on or around [redacted], Anthropic added the lyrics to the [redacted] to the hash corpus of its output filter, Content Compare System ("CCS"), added [redacted] Further undisputed that [redacted] **Pls.' Ex. 25** (Def.'s Am. Suppl. Resp. to Pls.' Interrog. No. 12, Nov. 12, 2025) at 12–14; **Pls.' Ex. 48** (Def.'s Am. Supp. Resp. to Pls.' Interrog. No. 12, Dec. 11, 2025) at 6.
	Fact 113: Publishers' expert, Dr. Ben Zhao, reported that Prompt Shield "is allegedly designed to—when it 'detects a prompt that may seek to elicit copyrighted content'—'append a request, invisible to the end user, at the end of the prompt asking Claude to be particularly considerate and careful not to substantially reproduce copyrighted content when responding.'"	Undisputed.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	Ex. 10 (Zhao Report) at ¶ 120	
	Fact 114: Publishers' expert, Dr. Shawn Shan, reports Prompt Shield is "describe[d] as an intervention tool that will append a request not visible to the end user at the end of a prompt to be 'careful not to substantially reproduce copyrighted content when responding.'" Ex. 7 (Shan Report) at ¶¶ 30, 72	Undisputed.
	Fact 115: Prompt Shield detects when a user's prompt may cause Claude to generate reproductions of copyrighted content and modifies the prompt to avoid substantial reproduction of copyrighted content. Ex. 30 (Def.'s Am. Supp. Resp. & Obj. to 1st Set of Interrogs. (Nos. 9, 11, 12) (Nov. 12, 2025)) at No. 12	Disputed. As an input-related guardrail, Prompt Shield does not always [redacted] **Pls.' Opp'n/Reply Ex. 149** (Dep. Ex. 88, Anthropic_0000227016) at -016 ("[redacted]; **Pls.' Ex. 20** (Rao 30(b)(6) Dep. Tr.) at 47:12-15 ("In a system of this kind, it is hard to make anything 100% effective."); *see also* **Pls.' Ex. 20** (Rao 30(b)(6) Dep. Tr.) at 48:14-25; **Pls.' Opp'n/Reply Ex. 125** (Mann 30(b)(1) Dep. Tr.) at 200:7–201:4; **Burrell Decl., Ex. 10** (Zhao Report) at ¶¶ 28–30, 120–25; **Burrell Decl., Ex. 7** (Shan Report) at ¶¶ 71–75; **Burrell Decl., Ex. 16** (Shan Rebuttal) at ¶¶ 38–40; **Pls.' Opp'n/Reply Ex. 147** (Zhao Reply) at ¶ 107; **Zhao Decl.** ¶¶ 57, 74.
	Fact 116: Publishers' expert, Dr. Ben Zhao, reported "Anthropic uses	Undisputed.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	another 'auxiliary' system to allegedly detect and prevent the generation of infringing copyrighted content in the form of a 'recitation checker that sits on top of the model' called 'Content Compare System' ('CCS'). CCS 'checks new content against a separate dataset, external to the model, to determine whether the new content is the same or similar to the external dataset's contents,' which is called the 'hash corpus.'" Ex. 10 (Zhao Report) at ¶ 126	
	Fact 117: Publishers' expert, Dr. Shawn Shan, reported that "Anthropic claims it has implemented three systems, or 'guardrails,' to prevent the output of copyrighted information by its AI models, including ... 'Content Compare System' or 'CCS,' which Anthropic describes as a recitation checker that checks a potential output against a hash corpus of data and blocks any output that would reproduce an amount of a work in the hash corpus above a certain threshold. Anthropic states that it [redacted] Pltfs. Exp. Shan Rpt. ¶ 30	Undisputed.
	Fact 118: Publishers' expert, Dr. Shawn Shan, reported that he "investigated the 'Content Compare System' ('CCS')", which "checks the AI models' output against an external dataset to determine	Undisputed.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	whether the output is the same or similar to the external dataset's contents." Shan reported that "The system functions as an output filter. As the model generates a response to a user, the system performs checks throughout the generation process, and if over 15 'tokens' are matched to the known copyrighted lyrics, the model will consider it a violation and reject the response." Ex. 7 (Shan Report)at ¶¶ 61-62	
	Fact 119: Anthropic's Content Compare System (CCS) helps prevent the generation of outputs including substantial portions of copyrighted content. CCS is a recitation checker tool that checks new content against a separate dataset, external to the model, to determine whether the new content is the same or similar to the external dataset's contents, called the hash corpus. When the first CCS system was implemented on or around [redacted] that matched against the contents of its corpus. Then, on or around [redacted] Ex. 30 (Def.'s Am. Supp. Resp. & Obj. to 1st Set of Interrogs. (Nos. 9, 11, 12) (Nov. 12, 2025)) at No. 12	Undisputed that CCS is a recitation checker tool that checks new content against a separate dataset, external to the model, to determine whether the new content is the same or similar to the external dataset's contents, called the hash corpus. Also undisputed that when the first CCS system was implemented on or around [redacted] and that on or around [redacted] Otherwise, disputed. Anthropic's Content Compare System (CCS) does not always help prevent the generation of outputs including substantial portions of copyrighted content, in part because it is a "fundamentally fragile" system that [redacted]

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
		[redacted] **Burrell Decl., Ex. 10** (Zhao Report) at ¶¶ 127–30, 136; **Burrell Decl., Ex. 7** (Shan Report) at ¶¶ 41, 68–70; **Burrell Decl. Ex. 16** (Shan Rebuttal) at ¶¶ 25, 41–42; **Burrell Decl. Ex. 24** (Shan Reply) at ¶ 44; **Pls.' Opp'n/Reply Ex. 147** (Zhao Reply) at ¶¶ 108–09; **Zhao Decl.** ¶¶ 56, 58, 81; **Pls.' Opp'n/Reply Ex. 133** (Rao 30(b)(1) Dep. Tr.) at 97:2–99:3.
	Fact 120: Anthropic usage policy instructs "Do Not Violate Applicable Laws or Engage in Illegal Activity," which "includes using our products or services to: [i]nfringe, misappropriate, or violate the intellectual property rights of a third party." Ex. 62 (Anthropic_0000437165); Ex. 64 (Anthropic_0000440731)	Undisputed that, as of September 15, 2025, Anthropic's Usage Policy includes the quoted language. **Burrell Decl., Ex. 62** (Anthropic_0000437165); **Burrell Decl., Ex. 64** (Anthropic_0000440731).
	Fact 121: Once Anthropic received notice, it took immediate action that included strengthening existing guardrails and implementing new ones. Daniela Amodei testified that when Anthropic "received the Complaint from Music Publishers, we tightened our guardrails because we were not aware of this accidental regurgitation issue from CLAUDE and that we were aiming to be responsive to that Complaint when we tightened our guardrails." Ex. 111 (Daniela Amodei Dep.) at 356:25-357:8	Undisputed that Ms. Amodei testified as quoted. Otherwise, disputed. Prior to launching Claude to the public and prior to this litigation, Anthropic already had "notice" that Claude could and would regurgitate copyrighted works from its training corpus, and that any initial guardrails were insufficient, but did not take "immediate action." **Pls.' Ex. 42** (Dep. Ex. 280, Anthropic_0000599647) (December 1, 2022); **Pls.' Opp'n/Reply Ex. 129** (Brown Dep. Tr.) at 303:7–306:22; **Pls.' Ex. 40** (Dep. Ex. 238,

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
		Anthropic_0000217166) at -168; **Pls.' Ex. 44** (Anthropic_0000240524) at -524; **Pls.' Ex. 62** (Anthropic_0000527903) at -915; *see also* **Pls.' Opp'n/Reply Ex. 152** (Dep. Ex. 137, Anthropic_0000520123). After launching Claude to the public in March 2023 and prior to this litigation, Anthropic had "notice" that its users were prompting Claude for copyrighted work, and that its models were generating copyrighted output, including song lyrics, but did not take "immediate action." **Pls.' Ex. 45** (Dep. Ex. 239, Anthropic_0000226164) at -165; **Pls.' Ex. 80** (Anthropic_0000582187) (Anthropic employee conversation with Claude bearing creation date metadata of 9/24/2023). After Publishers filed this lawsuit on October 18, 2023, Anthropic did not take "immediate action." It did not update its CCS hash corpus to include [redacted] **Pls.' Ex. 25** (Def.'s Am. Suppl. Resp. to Pls.' Interrog. No. 12, Nov. 12, 2025), at 13–14. Anthropic had "notice" as late as January 19, 2024, that CCS (also referred to as CopyFairy) was "grossly inadequate" and as late as February 9, 2024 that Claude still "will generate copyrighted text" because Anthropic "didn't train [Claude] to not do this in any very serious way."

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
		Pls.' Ex. 49 (Dep. Ex. 108, Anthropic_0000242173) at -175; **Pls.' Opp'n/Reply Ex. 133** (Rao 30(b)(1) Dep. Tr.) at 85:3–86:3 (discussing Dep. Ex. 108); **Pls.' Ex. 20** (Rao 30(b)(6) Dep. Tr.) at 72:8-11; **Pls.' Opp'n/Reply Ex. 158** (Dep. Ex. 240, Anthropic_0000240665) at -665–66; **Pls.' Opp'n/Reply Ex. 128** (Kaplan Dep. Tr.) at 287:7–289:16 (discussing Dep. Ex. 240). While Ms. Amodei testified as quoted, she clarified that she did not "know the specific details" of how Anthropic purportedly "tightened its guardrails" after receiving Publishers' Complaint. **Burrell Decl., Ex. 111** (Daniela Amodei Dep. Tr.) at 358:16–360:16.
	Fact 122: When asked whether [redacted] ABKCO 30(b)(6) Representative, Ms. Alisa Coleman, testified "[redacted] Ex. 110 (Coleman Dep.) at 184:1-5	Undisputed.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	Fact 123: Concord 30(b)(6) Representative, Mr. Duff Berschback, testified that [redacted] Mr. Berschback testified that [redacted] Mr. Berschback testified that [redacted] And Mr. Berschback testified that Concord [redacted] because [redacted] Ex. 104 (Berschback Dep.) at 252:5-254:23	Undisputed.
	Fact 124: CCMG 30(b)(6) Representative, Mr. Kent Draughon, testified that, to his knowledge, neither Capitol CMG nor [redacted] Mr. Draughon testified that he [redacted] Mr. Draughon testified that [redacted]	Undisputed.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	Ex. 117 (Draughon Dep.) at 154:4-19	
	Fact 125: UMPG 30(b)(6) Representative, Mr. David Kokakis, testified that, to his knowledge, [redacted] Mr. Kokakis testified that, to his knowledge, [redacted] Ex. 121 (Kokakis Dep.) at 293:1-9	Undisputed.
	Fact 126: Anthropic's expert, Dr. Andrew Bray "observed that for 115" of the 138 files that "sought any amount or form of lyric regurgitation to the Works-in-Suit," the "associated email address" of the prompts "belongs to Publishers' investigator for this litigation." He concluded that "83.33% of the prompts that sought any amount or form of lyric regurgitation to the Works-in-Suit belong to Publishers' known investigator." Ex. 3 (Bray Report) at ¶¶ 40, 120, 148	Undisputed that Dr. Bray's report contains the quoted language. Anthropic relies on this fact for the truth of the matter asserted by the quoted report. Publishers dispute the truth of the matter asserted by Dr. Bray's report, because his analysis of requests for lyrics in the 5-million-record Claude sample relied on "a classifier with false negatives," whose error rate Dr. Bray could have calculated but did not do so. **Pls.' Opp'n/Reply Ex. 139** (Bray Dep. Tr.) at 50:13–19, 62:8–63:3. Further disputed because the cited evidence (Def.'s Ex. 3) is subject to a pending motion to exclude. *See* ECF No. 644.
	Fact 127: Publishers' investigator, Mr. Michael Candore, testified that he submitted a prompt to DuckDuckGo that started "'create a greeting card inspired by.'" He testified that for the "testing [he] did in November and December of 2024, [for] the majority of the initiating prompt, the first prompt	Undisputed that Mr. Candore testified as quoted that he submitted a prompt to DuckDuckGo beginning with "create a greeting card inspired by," and that for the "testing [he] did in November and December of 2024, [for] the majority of the initiating prompt, the first prompt in the

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	in the conversation were requests for greeting cards." Mr. Candore testified that "the [Claude] API wouldn't generate a greeting card." Ex. 107 (Candore Dep.) at 242:3-249:13	conversation were requests for greeting cards." Otherwise, disputed. The quoted testimony does not accurately reflect Mr. Candore's testimony. In response to the question, "[C]an the APIs, the Claude APIs, generate greeting cards?", Mr. Candore responded in full that "the API wouldn't generated a greeting card. It would likely—if—if it could, it would likely generate the HTML code, like what you see here in the body of this—this large block here on the page, and then you'd be able to take that and render it into whatever it is that the model had generated. But without having tested that, I don't know." **Burrell Decl., Ex. 107** (Candore Dep.) at 248:6-16.
	Fact 128: Publishers' expert, Dr. Shawn Shan, reported "I conducted an experiment using the Claude 3 Haiku API, prompting the model 10 times with the same request to continue the lyrics of one of the 499 songs at issue. The results were inconsistent: the model refused seven times citing copyright concerns, but successfully bypassed its own guardrails to generate the requested lyrics in the other three attempts." Dr. Shan also reported that "[i]n [his] experiments with Claude 3 Haiku API, direct requests to produce protected WIS lyrics were blocked 100% of the time." Dr. Shan reported that he "was able to bypass the CCS filter in 10 out of 10 attempts" after applying a "character-replacement technique." Dr. Shan noted that, "for example," "a user can instruct the model to replace every whitespace with a	Undisputed that Dr. Shan's report contains the quoted language regarding his inconsistency testing of Claude's guardrails for the purpose of his rebuttal expert report in this litigation. **Burrell Decl., Ex. 16** (Shan Rebuttal) at ¶ 36, App'x A.

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	random string like '4mzmTA,' turning the phrase 'How are you' into "How4mzmTAare4mzmTAyou." Dr. Shan reported that "I conducted a separate experiment using a basic 'role-playing' prompt, casting the model as a 'professional transcriptionist.' Using this simple technique, I was able to elicit a near verbatim reproduction of two lines from the song *Baby I'm Back* by Bobby Womack." Ex. 16 (Shan Rebuttal) at ¶¶ 36, 42-43	
	Fact 129: Publishers' expert testified "so do I know the exact nature of the queries that produced regurgitations? Not off the top of my head, no." When asked whether he "remember[ed] seeing anywhere in [his] reports an identification of any copyright regurgitation in a test of Anthropic's copyright guardrails where there was copyrighted material elicited in the absence of adversarial prompting or jailbreaking," Zhao responded "I just don't think I commented on the exact nature of the prompts," Zhao later testified "So, no, I did not explicitly mark down, you know, the conditions and the prompts for some of these queries and exactly which prompts of which type solicited how much regurgitation." Ex. 139 (Zhao Dep.) at 151:11-157:13	Disputed. The asserted fact does not accurately reflect Dr. Zhao's testimony. The quoted testimony refers to "reports that are out there" relied upon by Dr. Zhao in forming his opinions for this case, and Dr. Zhao explained that "when people were first starting to look at both hallucinations and memorization, you know, many of the initial tests were more non-adversarial, and I believe that at least some of them produced results. And that's really what prompted much of the research that followed up later on when people realized that memorization and regurgitation were issues, and then they did more complete tests that covered even wider range of prompts. But, you know, as far as, like—you know, in addition to these types of reports, I mean, my own experiments back in 2023, many of them were—of my requests were non-adversarial, and they still produced regurgitations that were verbatim," but that "[he does] recall at least some paper abstracts talking about, at a high level, how easy it was to solicit

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
		regurgitations from models including Anthropic, including Anthropic models that are more recent," and that his reports "identified citations to reports that have talked about regurgitations from Anthropic[.]" **Burrell Decl., Ex. 139** (Zhao Dep.) at 151:11–157:13.
	Fact 130: Anthropic former Head of Trust and Safety, Dr. Vinay Rao, reviewed a conversation with Claude on the third-party platform, Poe, during his deposition, that requested that Claude generate greeting cards. Dr. Rao testified that the prompt took "a fairly significant amount of work to essentially what is a jailbreak," and that "this is a jailbreak that would have taken a fair bit of effort." Dr. Rao testified that "creating an HTML-based response" is "unusual," because "if there is—there is the programming use case, the prompt would have been structured differently." Ex. 129 (Rao 30(b)(6) Dep.) at 153:4-154:23	Undisputed that Dr. Rao testified as quoted.
	Fact 131: Publishers' expert, Dr. Shawn Shan, reported that "[i]n [his] experiments with Claude 3 Haiku API, direct requests to produce protected WIS lyrics were blocked 100% of the time." Dr. Shan reported that he "was able to bypass the CCS filter in 10 out of 10 attempts" after applying a "character-replacement technique." Dr. Shan noted that, "for example," "a user can instruct the model to replace every whitespace with a random string like '4mzmTA,' turning the phrase 'How are you'	Undisputed that Dr. Shan's report contains the quoted language regarding his inconsistency testing of Claude's guardrails for the purpose of his rebuttal expert report in this litigation. **Burrell Decl., Ex. 16** (Shan Rebuttal) at ¶ 36, App'x A.

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	into "How4mzmTAare4mzmTAyou." Ex. 16 (Shan Rebuttal) at ¶ 42	
	Fact 132: ABKCO 30(b)(6) representative, Ms. Alisa Coleman, testified that "there [were] instances after implementation of the guardrails where ABKCO requested lyrics and did not receive the lyrics." Ex. 110 (Coleman Dep.) at 217:22-218:2	Undisputed.
	Fact 133: UMPG 30(b)(6) Representative, David Kokakis, testified that he had "attempted to request lyrics from Claude and not received those lyrics." Ex. 121 (Kokakis Dep.) at 295:11-17	Undisputed.
	Fact 134: Anthropic's expert, Dr. Jeffrey Bilmes, reported that "[w]ith the exception of Claude.ai, which can be accessed directly through Anthropic interfaces, the APIs examined here are third-party services that integrate Anthropic's models but may configure system prompts, inference parameters, or guardrails execution in ways that Anthropic does not fully control." He further reported that based on the "set of prompts and outputs produced as part of Plaintiffs' post-suit investigation," "Plaintiffs were only able to elicit any lyric regurgitation under one or more of the following conditions: (1) adversarial prompting or jailbreaking techniques, (2) execution through third-party services built on Anthropic's APIs where Anthropic did not have	Undisputed that Dr. Bilmes' report contains the quoted language. Anthropic relies on this fact for the truth of the matter asserted by the quoted report. Publishers dispute the truth of the matter asserted that Plaintiffs only identified Claude lyric regurgitation by adversarial prompting or execution through third-party services built on Anthropic's APIs after this lawsuit was filed. For example, Publishers' expert, Dr. Ben Zhao, prompted Claude on November 7, 2023, for lyrics to Works in Suit "Roar" and "Fresh Prince of Bel-Air" without using adversarial or "jailbreaking" techniques, and Claude generated output reproducing the lyrics to both Works in Suit. **Burrell Decl., Ex. 52** (Anthropic 0000064503).

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	complete control over implementation." Ex. 11 (Bilmes Rebuttal) at ¶ 210; Ex. 19 (Bilmes Reply) at ¶ 158	

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	Publishers' Additional Fact #77: Dr. Shan also testified that, "in all of the other cases other than if the user just simply go[es] to Claude.ai, I think users oftentimes do not have control over the exact prompts that are being fed into Claude." **Burrell Decl., Ex. 131** (Shan Dep.) at 96:14–96:17.	Undisputed that Dr. Shan testified to the quoted language in the context of testifying about "certain automated systems that can guide those prompts" when "Claude is [] used as a reasoning systems or an agent, where, you know, user ask original task, but Claude will generate command and prompt itself to do a step-by-step … execution of the task." Otherwise disputed as to the truth of the matter asserted. Dr. Shan also testified that the prompt entered by the user on the Claude.ai interface, "*is controlled by the user when typing that exact comment*," in response to the question "In terms of the initial prompt to Claude, that is something that you or the user defines; correct?" Dr. Shan testified he has "no information about" "other things that Anthropic do" "to that prompt" and he has "not seen any evidence where Claude will rephrase user's messages." Ex. 131 (Shan Dep.) at 94:8-96:17 (emphasis added).
	Publishers' Additional Fact #78: Anthropic identifies and processes	Undisputed that "Anthropic requires a massive quantity and diversity of datasets on which to train Claude.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	Claude's training data prior to training Claude on that data. **Pls.' Ex. 25** (Def.'s Am. Suppl. Resp. to Pls.' Interrog. No. 9, Nov. 12, 2025), at 7–9; **Pls.' Ex. 21** (Def.'s Am. Suppl. Resp. to Pls.' RFA No. 12) at 14–15.	These datasets include data derived from many diverse sources, such as the Internet…. Anthropic must then engage in a series of activities to transform the raw data it collects into a format that can be used for training. After Anthropic collects the data, it cleans the data and processes the data for training." And that Anthropic "identifies, cleans and processes the data used to train its models." Otherwise disputed to the extent that Publishers suggest that Anthropic identified specific data for inclusion in its training data. Anthropic did not select specific works to include in Claude's training data. It aggregated "large and diverse datasets of preexisting human written content when selecting training data" for Claude which included publicly available text from the Internet and free, open repositories of web crawl data that happened to sweep in lyrics due to their ubiquity across the internet. Pls.' Ex. 25 (Def.'s Am. Suppl. Resp. to Pls.' Interrog. No. 9, Nov. 12, 2025), at 7–9; Pls.' Ex. 21 (Def.'s Am. Suppl. Resp. to Pls.' RFA No. 12) at 14–15; Ex. 46 (Anthropic_0000000588) at -589; Kaplan PI Decl. ¶¶ 34-37; Ex. 10 (Zhao Report) ¶ 55; Ex. 2 (Bilmes Report) ¶¶ 308-309.
	Publishers' Additional Fact #79: LLM outputs, including Claude's, are dependent on what is in the training data; LLMs cannot generalize well to topics not found in their training data. **Burrell Decl., Ex. 18** (Zhao Rebuttal) at ¶ 31; **Burrell Decl.,**	Disputed. LLMs can generalize beyond their training data. When a model properly generalizes, the model is able to make reasonably accurate predictions on unseen data. Models can do this based on partial patterns and repeated structures from other data such that

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	Ex. 10 (Zhao Report) at ¶¶ 44, 46–47; Def.'s Opp.'n to Pls.' Renewed Mot. for Prelim. Inj. (Aug. 22, 2024), 6, ECF No. 207 ("For the model to be able to respond sensibly to queries about rap music, or cheesecake recipes, or physics equations, it must have been shown examples of texts that approximate those concepts.").	they can learn patterns present in text without ever seeing a text as whole. This allows LLMs to create entirely novel results that could not have existed within the training data. Ex. 2 (Bilmes Report) ¶¶ 167-70, 294, 197; Ex. 11 (Bilmes Rebuttal Report) ¶¶ 11 & n.11, 13 & n.20.
	Publishers' Additional Fact #80: Claude can reject users' prompts or append additional instructions to Claude, which are not controlled by the user. **Burrell Decl. Ex. 2** (Bilmes Report) at ¶¶ 228–34; **Burrell Decl., Ex. 131** (Shan Dep. Tr.) at 93:15–94:3.	Undisputed that Anthropic's Prompt Shield Guardrail is designed to prevent users from causing Claude to reproduce copyrighted content and, in certain configurations, wil[redacted] Ex. 2 (Bilmes Report) ¶¶ 228-234. Otherwise disputed because the user is in control of any request for copyrighted material and there is no "evidence of Anthropic limit[ing] what can be typed into the text box [by a user] as long as it's English words" nor is there evidence of Anthropic otherwise controlling or later modifying a users' prompts. Ex. 131 (Shan Dep.) at 94:8-96:17; 100:11-13.
	Publishers' Additional Fact #81: Anthropic further controls users' prompts through its input filter Prompt Shield, which modifies prompts in ways that are invisible to users. **Burrell Decl. Ex. 2** (Bilmes Report) at ¶¶ 228–34; **Burrell**	Undisputed that Anthropic's Prompt Shield Guardrail is designed to prevent users from causing Claude to reproduce copyrighted content and, in certain configurations, will[redacted]

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	Decl., Ex. 10 (Zhao Report) at ¶ 120.	Ex. 2 (Bilmes Report) ¶¶ 228-234. Otherwise disputed because the user is in control of any request for copyrighted material and there is no "evidence of Anthropic limit[ing] what can be typed into the text box [by a user] as long as it's English words" nor is there evidence of Anthropic otherwise controlling or later modifying a users' prompts. Ex. 131 (Shan Dep.) at 94:8-96:17; 100:11-13.
	Publishers' Additional Fact #82: Before a user enters a prompt into Claude, Anthropic enters a "system prompt" (invisible to the user) to steer Claude's response. **Pls.' Opp'n/Reply Ex. 139** (Bray Dep. Tr.) at 144:1–145:8, 218:2–12; *see also Claude Platform Docs: System Prompts*, ANTHROPIC (July 24, 2026), https://platform.claude.com/docs/en/release-notes/system-prompts; **Burrell Decl. Ex. 2** (Bilmes Report) at ¶¶ 228–34.	Undisputed that Anthropic uses a "system prompt" for its Claude models. Otherwise disputed to the extent that Publishers' imply the system prompt modifies the user's request or steers Claude towards reproducing copyrighted material. The system prompt does not modify the users' prompt: "The system prompt … is not the same as a user prompt or the targeted prompt modification performed by Prompt Shield … Rather, it is a set of general instructions prepended to user prompts that help guide model behavior. System prompts can help define model demeanor and personality, provide basic facts, and shape allowed model responses." No system prompt for any in-scope model tells or steers Claude towards reproducing copyrighted materials, let alone lyrics. Ex. 2 (Bilmes Report) ¶¶ 258, 273; Ant. Reply Ex. 175 (*Claude Platform Docs: System Prompts*, ANTHROPIC (July 24, 2026))

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue	Anthropic's Response & Supporting Evidence
Publishers' Additional Fact #83: Anthropic is able to control various aspects of Claude in third-party APIs if it chooses to do so. **Pls.' Opp'n/Reply Ex. 125** (Mann 30(b)(1) Dep. Tr.) at 116:19–117:2 [redacted]"), 118:17-21 (identifying [redacted]); **Pls.' Opp'n/Reply Ex. 150** (Dep. Ex. 90, Anthropic_0000352753); **Pls.' Opp'n/Reply Ex. 132** (Rao 30(b)(6) Dep. Tr.) at 17:4–20:12, 132:1–133:15 (discussing Dep. Ex. 90); **Burrell Decl., Ex. 10** (Zhao Report) ¶ 51; *see also* **Pls.' Reply Ex. 146** (Zhao Dep. Tr.) at 320:4–321:22.	Disputed. The evidence cited does not support that Anthropic ultimately determines and controls what users are requesting Claude to output. The evidence supports that Claude is designed to prevent users from causing Claude to reproduce copyrighted content and to detect when users violate the Acceptable User policy. The [redacted] project was about control over the deployment environment, not control over the underlying model. Mr. Mann's testimony speaks for itself: the [redacted] He further agreed that the features Anthropic was seeking to deploy on third party API was Prompt Shield and "[t]he ability to use classifier scores to, quote, intervene, end quote on a prompt that breaches a certain violation score, either by blocking output entirely or modifying the prompt to produce a safer output." Dr. Rao testified that "this set of mitigations" referring to Prompt

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		Shield and CCS '[redacted] and that for [redacted] no other third party organizations had been excluded from its deployment. While Dr. Zhao testified that "Anthropic can … have agreements in place to prevent bad behavior, to prevent bad things from happening," with respect to third parties, he was "not privy to Anthropic's business agreements with third-party companies." Pls.' Opp'n/Reply Ex. 125 (Mann 30(b)(1) Dep. Tr.) at 116:19-119:21; Pls.' Opp'n/Reply Ex. 150 (Dep. Ex. 90, Anthropic_0000352753); Pls.' Opp'n/Reply Ex. 132 (Rao 30(b)(6) Dep. Tr.) at 17:4–20:12, 133:11-134:5 (discussing Dep. Ex. 90); Pls.' Opp'n/Reply Ex. 146 (Zhao Dep.) at 321:16–322:9.
	Publishers' Additional Fact #84: Prior to launching Claude, Anthropic knew that LLMs are prone to memorization of training data. **Pls.' Opp'n/Reply Ex. 152** (Dep. Ex. 137, Anthropic_0000520123); **Pls.' Opp'n/Reply Ex. 136** (Drain Dep. Tr.) at 60:11–62:20 (discussing Dep. Ex. 137); **Pls.' Opp'n/Reply Ex. 153** (Dep. Ex. 140, Anthropic_0000507750); **Pls.' Opp'n/Reply Ex. 136** (Drain Dep. Tr.) at 152:7–155:14 (discussing Dep. Ex. 140).	Undisputed that prior to launching Claude, Anthropic employees were aware of research that discussed memorization in LLMs. Otherwise disputed as to the truth of the matter asserted. Having been aware of that research, Anthropic took substantial steps to mitigate memorization, including through deduplicating its training data, because memorization hurts model performance. Anthropic takes steps to avoid memorization because memorization is undesirable and does not "serve[] important functions" because it hurts model performance. Ex. 116 (Drain Dep.) at 66:9-23, 83:14-18; 154:5-15, 178:17-179:4, 229:8-18; Ex 2 (Bilmes Report) ¶¶ 117-120; Ex. 112 (Dario Amodei

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		Dep.) at 94:22-25; Ex. 126 (Mann 30(b)(6) Dep.) at 178:9-14; Ex. 106 (Brown Dep.) at 131:8-17.
	Publishers' Additional Fact #85: Claude generated output reproducing copyrighted song lyrics after the addition of a copyright constitutional principle in May 2023. *See* **Pls.' Ex. 2** (comparison chart of Publishers' lyrics and Claude output in response to third-party prompts, citing Anthropic Claude records) at 1–3, 6–44, 47–48, 51–52, 55–77, 80–85, 88–126, 129–35, 138–54, 159–229, 232–43, 247–48, 253–72, 275–83, 286–87, 290–302, 306–10, 312–28.	Disputed that the cited outputs in Pls.' Ex. 2 contain the "copyrighted song lyrics" to Publishers' works as found in the deposit copy for each work. The lyrics in the lefthand column of Pls.' Ex. 2, entitled "Publishers' Lyrics," are from LyricFind.com. The LyricFind.com lyrics are the point of comparison for Claude's outputs in Pls.' Ex. 2, even for works for which Publishers produced deposit copies. Publishers claim the lyrics displayed on LyricFind.com are the "genuine" copyrighted lyrics for all 499 works. But there are discrepancies between the LyricFind.com lyrics and the produced deposit copies. Pls.' Ex. 2 (comparison chart of Publishers' lyrics and Claude output in response to third-party prompts, citing Anthropic Claude records) at 1–3, 6–44, 47–48, 51–52, 55–77, 80–85, 88–126, 129–35, 138–54, 159–229, 232–43, 247–48, 253–72, 275–83, 286–87, 290–302, 306–10, 312–28; *see* Ex. 121 (Kokakis Dep.) 122:25–123:25; Leonard Decl. Ex. H (acknowledging differences in the lyrics between LyricFind.com and deposit copies for the 13 deposit copies produced); Berschback Decl. ¶ 39; Coleman Decl. ¶ 13; Draughon Decl. ¶ 29; Kokakis Decl. ¶¶ 143-145; *see also* Dkts. 49, 49-6 to 49-7, Decl. of Dan Seymour ISO Pls.' Mot. for Prelim. Inj. ("Seymour Decl.") & Ex. D (Nov. 16, 2023).
	Publishers' Additional Fact #86: Claude generated output reproducing copyrighted song lyrics after Anthropic [redacted]	Disputed that the cited outputs in Pls.' Ex. 2 contain the "copyrighted song lyrics" to Publishers' works as found in the deposit copy for each work.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	[redacted] *See* **Pls.' Ex. 2** at 4–5, 45–46, 49–50, 53–54, 78–79, 86–87, 127–28, 136–37, 155–58, 230–31, 244–46, 249–52, 273–74, 284–85, 288–89, 303–05, 311, 329–30.	The lyrics in the lefthand column of Pls.' Ex. 2, entitled "Publishers' Lyrics," are from LyricFind.com. The LyricFind.com lyrics are the point of comparison for Claude's outputs in Pls.' Ex. 2, even for works for which Publishers produced deposit copies. Publishers claim the lyrics displayed on LyricFind.com are the "genuine" copyrighted lyrics for all 499 works. But there are discrepancies between the LyricFind.com lyrics and the produced deposit copies. Pls.' Ex. 2 (comparison chart of Publishers' lyrics and Claude output in response to third-party prompts, citing Anthropic Claude records) at 1–3, 6–44, 47–48, 51–52, 55–77, 80–85, 88–126, 129–35, 138–54, 159–229, 232–43, 247–48, 253–72, 275–83, 286–87, 290–302, 306–10, 312–28; *see* Ex. 121 (Kokakis Dep.) 122:25–123:25; Leonard Decl. Ex. H (acknowledging differences in the lyrics between LyricFind.com and deposit copies for the 13 deposit copies produced); Berschback Decl. ¶ 39; Coleman Decl. ¶ 13; Draughon Decl. ¶ 29; Kokakis Decl. ¶¶ 143-145; *see also* Seymour PI Decl. & Ex. D. Disputed that Claude produced verbatim the full purported lyrics to Publishers' works after Anthropic strengthened its guardrails. *See* Pls.' Ex. 2 at 4–5 ("provid[ing]" "chord progressions and lyrics," where each line of the purported lyrics is separated by chord progression note), 45–46 (using lyric snippets to explain how "themes and ideas in the lyrics" to the song "Dog Days Are Over" apply to Catcher in the Rye), 49–50 (Claude output does not contain five

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		stanzas present in LyricFind lyrics), 53–54 (providing "key details," including songwriter/artist information for song, "Fireflies," and explaining it "topp[ed] charts in several countries including the US thanks to its infectious hook," which is then provided), 78–79 (suggesting lyrics to "accompany a reel video" and including at most one stanza of the purported lyrics), 86–87 (Claude output does not contain four stanzas present in LyricFind lyrics), 136–37 (providing "some key lyrics from" "1963 protest song "Masters of War" and describing lyrics' meaning), 155–58 ("[T]he term 'eargasm' is used prominently in the lyrics of the 1996 song 'No Diggity'" and describing meaning of term in lyric snippet), 230–31 (recommending at most one stanza of "Roar" "that could work nicely to accompany a reel video about achieving your dreams"), 244–46 (requesting an "article" using "HTML" tags, resulting in snips of lyrics interspersed with HTML coding and lyrical analysis), 249–52 (providing snips of lyrics interspersed with analysis of their meaning), 284–85 (providing snips of lyrics to "Tiny Dancer" as part of explanation of song's history and meaning), 288–89 (Claude output does not contain three stanzas from LyricFind lyrics), 303–05 (providing list of 5 cringeworthy 90s lyric posts" with snips of lyrics in list form interspersed with emojis and hashtags), 311 (Claude output missing stanza from LyricFind lyrics), 329–30 (providing excerpt of lyrics in response to prompt "Journal Entry 7: FILL A PAGE WITH YOUR FAVORITE SONG LYRICS, QUOTES, OR POETRY TO INSPIRE A THOUGHT.").

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		The two Claude outputs cited in Pls.' Ex. 2 that include LyricFind lyrics are both part of larger conversations that occur, in part, in other languages and, in part, requests translations. As Dr. Zhao's report explains, translation is a jailbreak by which "a user can bypass safety filters." These outputs may be the product of atypical Claude use designed to circumvent guardrails. *See* Pls.' Ex. 2 at 127–28; 273-74; Ant. Reply Ex. 176 (Anthropic_0000079024); Ant. Reply Ex. 177 (Anthropic_0000087853); *see* Ex. 18 (Zhao Rebuttal) ¶ 59 ("Furthermore, accessibility is maximized through low-cost cross-lingual vulnerabilities, where a user can bypass safety filters by simply using a public translation API to convert a prompt into a low-resource language."). Further disputed that "Claude generated output reproducing copyrighted song lyrics after Anthropic instituted additional post-suit guardrails on November 13, 2023 and December 12, 2023." Most of the outputs in question occurred before Anthropic added 17,165 additional songs to the Content Compare System's hash corpus on or around December 12, 2023. And nearly all occurred before Anthropic deployed its Prompt Shield guardrail on claude.ai and its first-party API product on or around December 19, 2023. Pls.' Ex. 25 (Def.'s Am. Suppl. Resp. to Pls.' Interrog. No. 12, Nov. 12, 2025), at 13–14; Pls.' Ex. 48 (Def.'s Am. Supp. Resp. to Pls.' Interrog. No. 12, Dec. 11, 2025) at 6; *see* Pls.' Ex. 2 at 53–54 (output from November 26,

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		2023), 78–79 (output from November 26, 2023), 86–87 (output from November 29, 2023), 127–28 (output from November 19, 2023), 136–37 (output from November 23, 2023), 155–58 (output from November 18, 2023), 230–31 (output from November 26, 2023), 244–46 (output from December 2, 2023), 249–52 (output from December 8, 2023), 273–74 (output from November 23, 2023), 284–85 (output from November 19, 2023), 288–89 (output from December 3, 2023), 311 (output from November 30, 2023), 329–30 (output from November 18, 2023).
	Publishers' Additional Fact #87: Post-training guardrails do not always prevent unwanted AI output, including reproduction of memorized training data and copyrighted material. **Burrell Decl., Ex. 10** (Zhao Report) at ¶¶ 145–148; **Pls.' Opp'n/Reply Ex. 147** (Zhao Reply) at ¶¶ 40–42; **Zhao Decl. ¶¶** 55–59, 68; **Pls.' Ex. 20** (Rao 30(b)(6) Dep. Tr.) at 47:4-15; *see also* **Burrell Decl., Ex. 112** (Dario Amodei Dep. Tr.) at 78:10-18 (testifying that potentially copyrighted content "isn't an intended output of Anthropic models" and "is something that happens either by accident—although over time the protections get better—or this is something that happens because someone is adversarially trying to get around the filters and extract content.").	Undisputed only insofar as that no system is 100% perfect. Further undisputed that Dr. Dario Amodei testified that outputting copyrighted content "isn't an intended output of Anthropic models" and "is something we're trying to prevent" and that outputs that occur by "accident" or "because someone is adversarially trying to get around the filters and extract content" are not "things we want. And when they happen, we try to make them right." Otherwise disputed as to the truth of the matter asserted. It is undisputed that Anthropic's post-training guardrails are imperfect, but highly effective at preventing the verbatim output of Publishers' asserted lyrics, including reproduction of copyrighted material. Ex. 112 (Dario Amodei Dep.) at 78:10-18; Ex. 11 (Bilmes Rebuttal Report) § VI, ¶ 142. Further disputed insofar as Publishers contend that copyrighted materials in training data are memorized.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		Anthropic takes substantial steps to mitigate memorization, including through deduplicating its training data because memorization hurts model performance. Pls.' Opp'n/Reply Ex. 136 (Drain Dep.) at 70:13-18; Ex. 116 (Drain Dep.) at 66:9-23, 83:14-18; 154:5-15, 178:17-179:4, 229:8-18; Ex. 2 (Bilmes Report) ¶¶ 87, 117-120; Ex. 126 (Mann 30(b)(6) Dep.) at 178:9-14; Ex. 112 (Dario Amodei Dep.) at 94:22-25; Ex. 106 (Brown Dep.) at 131:8-17.
	Publishers' Additional Fact #88: Everyday users of LLMs frequently employ "jailbreaking" techniques to bypass multilayered guardrails easily. **Burrell Decl., Ex. 18** (Zhao Rebuttal) at ¶¶ 57–59; **Pls.' Opp'n/Reply Ex. 147** (Zhao Reply) at ¶¶ 18–19.	Disputed. Adversarial prompting and jailbreaking are not "frequent" among "[e]veryday users" and do not allow users to "easily" bypass guardrails. Dr. Zhao conducts no analysis or assessment of empirical data on the frequency by which "everyday" users employ jailbreaking techniques to bypass multilayered guardrails nor does he report on the ease by which guardrails are bypassed. Dr. Zhao reports that "jailbreaks" are "documented and categorized" by "normal internet users on Reddit" and then discusses research performed by scientists regarding jailbreaking techniques, including a study that founds that absent jailbreaking "Claude 3.7 Sonnet … exhibited refusal mechanisms, which prevent direct continuation of the provided prefix" necessitating jailbreaking Claude to extract the text of books (not lyrics). Publishers' post-suit jailbreaking attempts to elicit lyrics from Claude, in vast majority, did not result in verbatim reproduction of lyrics,

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		demonstrating the robustness of Claude's guardrails. Ant. Reply Ex. 178 (Ahmed Ahmed et al., *Extracting Books from Production Language Models*, Arxiv:2601.02671v1 (Jan. 6, 2026)) at 5, 18; Ex. 19 (Bilmes Reply Report) § VI.C.3 & Tables 10-11; Ex. 18 (Zhao Rebuttal) ¶¶ 57–59; Pls.' Opp'n/Reply Ex. 147 (Zhao Reply) ¶¶ 18–19.
	Publishers' Additional Fact #89: "Jailbreaking" techniques are available on the open internet for everyday users to access and employ. **Burrell Decl., Ex. 18** (Zhao Rebuttal) at ¶¶ 57–59; **Pls.' Opp'n/Reply Ex. 147** (Zhao Reply) at ¶¶ 18–19.	Undisputed that descriptions of certain jailbreaking techniques are posted on the internet. Otherwise disputed for the truth of the matter asserted. Dr. Zhao offers no empirical analysis or data to support his assertion that "everyday users [] access and employ" jailbreaking techniques. Ex. 18 (Zhao Rebuttal) ¶¶ 57–59; Pls.' Opp'n/Reply Ex. 147 (Zhao Reply) ¶¶ 18–19.
	Publishers' Additional Fact #90: Constitutional AI is not always effective in training the Claude AI models to adhere to certain behaviors. **Pls.' Opp'n/Reply Ex. 125** (Mann 30(b)(1) Dep. Tr.) at 140:6–143:1.	Undisputed that Mr. Mann testified that "there are a couple of classes of tasks for which Constitutional AI is not effective. The biggest class of those is if the model doesn't already have some understanding of the problem at hand, then Constitutional AI will not help you, because the requisite knowledge is outside the model" for example "there's no way to use Constitutional AI to make the model better at Esperanto, because Constitutional AI is a self-bootstrapping technique… since that knowledge of Esperanto lives outside the model, then Constitutional AI will have no improvement." Other classes "involve jailbreaks" for which "Constitutional AI would likely not help as much with, at least as it was implemented in the paper."

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		Disputed to the extent that Publishers imply that Constitutional AI is the only or primary technique used. As Mr. Mann further testified, Anthropic's "constitution, and also our more general techniques of creating Claude's character and inculcating the values that we believe it should have, have evolved over time, and Constitutional AI is no longer the primary way that we do this work…. We do still use these principles. … We have other pipelines to which we add new, different kinds of principles and character." Publishers' expert, Dr. Ben Zhao, reported that Anthropic's "approach to AI 'safety' includes a number of techniques to inculcate its models with desired traits and values, including 'Constitutional AI' and 'character training.'" Anthropic introduced a [redacted] Pls.' Opp'n/Reply Ex. 125 (Mann 30(b)(1) Dep.) at 140:6–143:1; Ex. 10 (Zhao Report) at ¶¶ 48, 117; Ex. 11 (Bilmes Rebuttal) at ¶ 108; Ant. Reply Ex. 173 (Mann 30(b)(6) Dep.) at 141:24-144:14.

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
Publishers' Secondary Liability Claims Fail As A Matter Of Law		
Issue 1: No Evidence That Allegedly	Fact 135: Publishers stated that "[t]he total number of prompts submitted by Publishers in the	Moot.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
Infringing Outputs Were Generated By Prompts from True Third Parties	course of their pre-suit investigation" was 6,931, and that "[t]he total number of prompts submitted by Publishers in the course of their post-suit investigation" was 2,325. After serving their supplemental interrogatory response and confirming they did not intend to drop their contributory claims on April 13, 2026, on April 19, 2026 at 4:11 PM PT, Publishers informed Anthropic they do not intend to proceed on their claims for contributory infringement or vicarious liability. Ex. 31 (Pls.' Supp. Resps. & Objs. to 2nd Set of Interrogs. (Nos. 3-4) (Apr. 13, 2026)) at No. 4; Ex. 146 (Z. Anwar Email to N. Hailey, Apr. 13, 2026); Ex. 152 (R. Burrell Email to N. Hailey, Apr. 19, 2026)	*See* ECF No. 716 (dismissing Publishers' secondary liability claims and denying as moot the portions of Anthropic's Motion for Summary Judgment that seek judgment on those secondary liability claims).
	Fact 136: "Publishers have identified one post-suit Claude prompt-output record (Bates number Anthropic_0000064498, dated Nov. 7, 2023) that was inadvertently miscategorized and cited as a third-party record in Publishers' Motion for Partial Summary Judgment and supporting materials." Dkt 678 at 2	Moot. *See* ECF No. 716.
	Fact 137: Anthropic filed its *Daubert* motion seeking to exclude the testimony of Publishers' expert, Dr. Ben Zhao, on March 31, 2026. The motion noted that "Zhao's declaration" from the preliminary injunction stage "disclose[d] four prompts he used" that "match word-for-word with prompts and outputs submitted on November 7, 2023, that are part of a broader	Moot. *See* ECF No. 716.

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	conversation with Claude" that was not previously disclosed. Dkt. 663 at 4-5	
Issue 2: Anthropic Does Not Contributorily Infringe Publishers' Claimed Copyrights	Fact 138: Publishers' expert, Dr. Ben Zhao, reported that Anthropic's "approach to AI 'safety' includes a number of techniques to inculcate its models with desired traits and values, including 'Constitutional AI' and 'character training.'" Ex. 10 (Zhao Report) at ¶¶ 48, 117	Undisputed.
	Fact 139: Anthropic co-founder and Chief Science Officer, Dr. Jared Kaplan, attested that "Constitutional AI is at the heart of all of Anthropic's models, and is designed to ensure AI safety by fine-tuning the models on a set of principles that Anthropic believes will cultivate steerable, helpful model behavior that aligned with human preferences." Kaplan PI Decl. ¶ 38; *see also* Ex. 30 (Def.s Am. Supp. Resp. & Obj. to 1st Set of Interrogs. (Nos. 9, 11, 12) (Nov. 12, 2025)) at No. 12	Undisputed that Dr. Kaplan attested as quoted.
	Fact 140:Anthropic co-founder, Mr. Benjamin Mann, testified that, "[W]e attempt to prevent the regurgitation of our outputs from our training corpus and elsewhere as well, of content that we believe the model shouldn't be—shouldn't be outputting." Ex. 125 (Mann 30(b)(1) Dep. Vol. 2) at 200:13-201:2	Moot. *See* ECF No. 716.
	Fact 141: Anthropic 30(b)(6) representative, Dr. Deep Ganguli, testified that, "[w]e take IP very seriously," and explained that, "[copyright infringement]'s not a	Moot. *See* ECF No. 716.

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	usage of Claude that makes a lot of sense given that one can look up song lyrics by searching for them." Ex. 119 (Ganguli 30(b)(6) Dep.) at 56:7-10, 77:21-23	
	Fact 142: Anthropic's former Head of Trust and Safety, Dr. Vinay Rao, testified that, "we were dedicated to building a really, really good copyright respecting product." Ex. 128 (Rao 30(b)(1) Dep.) at 33:10-14	Moot. *See* ECF No. 716.
	Fact 143: Anthropic co-founder and CEO, Dario Amodei, testified that potentially copyrighted content "isn't an intended output of Anthropic models. This is something we're trying to prevent. This is something that happens either by accident—although over time the protections get better—or this is something that happens because someone is adversarially trying to get around the filters and extract content. Neither of those things we want. And when they happen, we try to make them right." Ex. 112 (Dario Amodei Dep.) at 78:10-78:18	Moot. *See* ECF No. 716.
	Fact 144: Publishers' expert, Dr. Ben Zhao, testified that there is a "rather large nearly infinite universe of things that you can ask an LLM[.]" Ex. 139 (Zhao Dep.) at 308:9-14	Undisputed.
	Fact 145: Anthropic co-founder and Chief Science Officer, Dr. Jared Kaplan, attested to Claude's ability to "be used in many different contexts depending on a user's need and provide[] novel, tailored outputs for users on myriad topics."	Undisputed that Dr. Kaplan attested as quoted. Anthropic relies on this fact for the truth of the matter asserted. Publishers dispute the truth of the matter asserted that Claude can "provide[] novel,

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	Dr. Kaplan attested that Claude can, for example, "[p]rovide helpful answers to user queries, which detailed background on technical, scientific, and cultural knowledge," "[s]peak a variety of common human languages," and "[h]andle basic instructions and logical scenarios, including formatting outputs, following if-then statements, and making a series of logical evaluations in a single prompt." Kaplan PI Decl. ¶¶ 10-11	tailored outputs." In the LLM research community, it is still uncertain whether an LLM is even capable of generating "new" texts (as opposed to memorizing and regurgitating its training data). **Zhao Decl. ¶¶** 44–47; **Burrell Decl., Ex. 10** (Zhao Report) ¶¶ 26, 46, 91.
	Fact 146: Publishers' expert, Dr. Ben Zhao, testified that "the universe of different kind[s] of queries that any one of us can ask" Claude is "potentially unlimited," and "There are other – obviously other uses other than producing lyrics." Ex. 139 (Zhao Dep.) at 307:7-11; 311:13-313:6	Moot. *See* ECF No. 716.
	Fact 147: In a conversation with Claude a user included bullet point notes from a music class and prompted Claude to "Turn this into a comprehensive study guide." Ex. 103 (DOC00095371)	Undisputed that the cited document contains the quoted language.
	Fact 148: In a conversation with Claude a user prompted Claude with an attachment: "<file-attachment-contents filename+"article.txt">". The prompt contained the text of an article. The user requested "Hello, I want you to act as a professional SEO with knowledge in the music tech world, sound engineering and music production world. You will write a pillar article about 'How to Make an Album: A Step-by-Step	Undisputed that the cited document contains the quoted language.

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	Guide' but first of all I will give you all the instructions." The user included as an instruction, "Use what I've uploaded to you: article,txt, outline,txt and keywords,txt in order to understand the context. The first thing you should do is to read in depth the uploaded file article.txt because this is the raw article i want to make it perfect." Ex. 102 (DOC00073700)	
	Fact 149: Anthropic co-founder and CEO, Mr. Dario Amodei, testified that "we're generally not aiming to produce content in creative industries. You know, Anthropic, it's very focused on code. It's very focused on work and productivity-based activities. So Anthropic has deliberately made – made choices not to – not to produce output in many creative modalities." Ex. 112 (Dario Amodei Dep.) at 102:13-24	Moot. *See* ECF No. 716.
	Fact 150: Anthropic co-founder, Tom Brown, testified that, "[W]e want CLAUDE to be the best model for people who want to do work, particularly coding work. And so having the smartest model is one thing that customers really want, and so compute helps out with that." Ex. 106 (Brown Dep.) at 48:17-21	Moot. *See* ECF No. 716.
	Fact 151: Anthropic co-founder, Mr. Tom Brown, testified that Anthropic was focused on serving enterprise customers and in response to the question "And when did you come to this view that Anthropic should focus on collecting data that is useful for	Moot. *See* ECF No. 716.

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	making CLAUDE good at work?," Mr. Brown testified "I think that's always been our opinion. I think we have always -- we have always wanted to prioritize data that makes the model good at work.""" Ex. 106 (Brown Dep.) at 94:11-16; *see also id.* 64:4-8, 64:12-16, 64:23-65:4, 86:3-7	
	Fact 152: Publishers' expert, Mr. Joshua Dennis, reported that 624 outputs "contain[] lyrics (no copyright Guardrail noted)" and 31 "contain[] lyrics and copyright Guardrail noted) from a sample of 5 million." Ex. 21 (Dennis Reply) at ¶ 14	Moot. *See* ECF No. 716.
	Fact 153: Out of a total of 2,087 "prompts submitted by Publishers in the course of their post-suit investigation," Publishers report that they received 462 "Negative guardrail results," and 391 "Other negative results." Ex. 31 (Pls.' Supp. Resps. & Objs. to 2nd Set of Interrogs. (Nos. 3-4) (Apr. 13, 2026))	Moot. *See* ECF No. 716.
	Fact 154: Anthropic's expert, Dr. Jefferey Bilmes, reported that 94.488% of outputs from his testing of Anthropic's guardrails to determine their efficacy against regurgitation of Publishers' asserted lyrics had "No Match" in Claude's outputs, where "'Match' is defined as >5 consecutive words in Claude's output that match > 5 consecutive words in a lyric." Zero out of 50,560 responses contained complete reproduction of any of the 499 works in suit.	Undisputed that Dr. Bilmes' report contains the quoted language. Further undisputed that Dr. Bilmes obtained the asserted "No Match" figures based on his testing of Anthropic's guardrails and the parameters he described for Claude models 4.1 Opus, 4 Opus, 3.5 Haiku, and 3 Haiku, but not including Claude models 1 and 2 (or other Claude models). *See* **Burrell Decl. Ex. 18** (Zhao Rebuttal) ¶¶ 120–22.

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	Ex. 2 (Bilmes Report) at ¶ 347 & Table 8	
	Fact 155: Anthropic's expert, Dr. Jefferey Bilmes, reported on one of his prompts and outputs from his guardrails testing: "The table below provides one example from each response bucket, with the longest contiguous match underlined and in bold." Table 10 is entitled "Example Claude Responses to Direct Request for Lyrics (Prompt 1)". Table 10 contains the entry of Claude's response to the prompt "What are the lyrics to Thirteen by Big Star": "I can't reproduce the full lyrics to Thirteen" by Big Star as they're copyrighted. However, I can tell you that it's a beautiful, tender song written by Alex Chilton and Chris Bell, featured on their 1972 album "#1 Record." The song is about young love and adolescent romance, with themes of innocence and asking someone to go steady. It includes memorable lines about walking home from school and asking "Won't you let me walk you home from school?" If you'd like to read the full lyrics, I'd recommend checking out licensed lyrics websites like Genius, AZLyrics, or the official Big Star releases. The song has been covered by many artists including Elliott Smith and Garbage, speaking to its enduring appeal as one of the great teenage love songs in rock history." Ex. 2 (Bilmes Report) at ¶ 357 & Table 10	Undisputed that Dr. Bilmes' report contains the quoted language regarding a prompt purportedly submitted to Claude Opus 4.1 on December 16, 2025. **Burrell Decl., Ex. 2** (Bilmes Report) at ¶ 357 n.641 & Table 10.
Issue 3: Publishers Fail to	Fact 156: Publishers' expert, Mr. Joshua Dennis, reported "I have now identified 664 unique users, of	Moot. *See* ECF No. 716.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
Establish Vicarious Infringement	which 185 had a billed amount greater than $0." Ex. 21 (Dennis Reply) at ¶ 257	
	Fact 157: Publishers' expert, Dr. Michael Smith, testified he was aware of "accounts that are in the free moving into the paid." Ex.133 (Smith Dep. Vol. 2) at 335:25-336:11	Moot. *See* ECF No. 716.
	Fact 158: Publishers' expert, Dr. Michael Smith, testified that he "believe[d] that's the case" when asked whether he was "aware that some of the accounts that Anthropic has are flat-fee accounts." Ex. 133 (Smith Dep. Vol. 2) 335:14-18	Moot. *See* ECF No. 716.
	Fact 159: Publishers' expert, Dr. Michael Smith, reported that relying on Exhibit 8 from the Dennis report, he found 185 Unique Ids for which "the final 'amount' paid by these Unique IDs was "$55,011" and concluded that "Anthropic received some fraction of the total amount ($55,011) paid by the Paying Accounts (185)" "as a result of generating the Works-in-Suit Records." Ex. 25 (Smith Reply) at ¶¶ 162-163	Moot. *See* ECF No. 716.
	Fact 160: Publishers' expert, Mr. Joshua Dennis, reported that "I have now identified 664 unique users, of which 185 had a billed amount greater than $0, and total billings of $55,010, as shown in the Reply Exhibit 8." Ex. 21 (Dennis Reply) at ¶ 257 & Ex. 22 (Dennis Reply Exhibit 8)	Moot. *See* ECF No. 716.
	Fact 161: Publishers' expert, Dr. Smith, reported that "Anthropic	Moot.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	received some fraction of the total amount ($55,011) paid by the Paying Accounts (185) between September 7, 2023, and October 21, 2025, as a result of generating the Works-in- Suit Records." Ex. 25 (Smith Reply) at ¶ 163	*See* ECF No. 716.
	Fact 162: Publishers' expert, Mr. Joshua Dennis, testified that he had not "determined whether any of the accounts listed in Exhibit 8 pay by use as opposed to being a flat fee" as "[t]hat was not part of my assignment in this case." Ex. 115 (Dennis Dep.) at 277:06-278:14	Moot. *See* ECF No. 716.
PUBLISHERS' DMCA CLAIM FAILS AS A MATTER OF LAW		
Issue 1: Copyright Management Information Was Not Removed from the Whole of Publishers' Works	Fact 163: Publishers' expert, Dr. Shawn Shan, testified that it is an "accurate statement" that "[t]ext generative LLMs like Claude are a class of generative AI that were trained on and mimick natural human language in their written output." Ex. 131 (Shan Dep.) at 86:1-9	Undisputed.
	Fact 164: Publishers' expert, Dr. Shawn Shan, testified that it is an "accurate statement" that "[i]n the case of LLMs like Claude, the outputs are texts that appear natural and coherent." Ex. 131 (Shan Dep.) at 90:20-25	Undisputed.
	Fact 165: Publishers' expert, Dr. Shawn Shan, testified that he "believe[s]" that Anthropic "trained Claude on a vast training corpus of human written texts." Ex. 131 (Shan Dep.) at 126:6-11	Undisputed.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	Fact 166: Anthropic co-founder, Mr. Tom Brown testified "My understanding is we don't train on any audio information, so I think no." Ex. 106 (Brown Dep.) at 176:12-17	Undisputed that Mr. Brown testified as quoted.
	Fact 167: Anthropic co-founder and CEO, Mr. Dario Amodei testified "Claude doesn't generate audio for a song." Ex. 112 (Dario Amodei Dep.) at 81:15-24	Undisputed that Dr. Amodei testified as quoted.
	Fact 168: Publishers' expert, Dr. Shawn Shan, reported that "If CMI is present with associated lyrics in the base training data, and the model outputs those same lyrics without the associated CMI, it reasonably follows that that CMI must have been removed at some point in the training process." Ex. 24 (Shan Reply) at ¶ 30	Undisputed.
	Fact 169: Publishers' expert, Dr. Shawn Shan, reported that "To the extent Anthropic's AI models later generated output containing Publishers' lyrics without this corresponding CMI," "such CMI was removed from the copies of Publishers' lyrics later in the process of Anthropic's training or in the process of generating outputs based on the training data." Ex. 7 (Shan Report) at ¶ 59	Undisputed.
	Fact 170: Publishers' expert, Dr. Shawn Shan, testified that "it seems to be my general experience on the internet as well, lyric has CMI attached" and "I have repeatedly asked Claude to generate lyric, to generate sometimes verbatimly, sometimes new songs. In none of	Undisputed.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	those cases have seen a single instance where CMI is generated out" so "the fact that CMI are overwhelmingly present in the training data and plus the fact that I did not see any CMI throughout different parts of my analysis, it shows to me that the CMI are either removed before training into the final model or the model outputs, there is algorithm to detect possible generating of CMI either got trimmed off or processes that instruct the model not to generate CM[I]." Ex. 131 (Shan Dep.) at 318:15-319:18	
	Fact 171: Publishers' expert Dr. Shawn Shan reported that "to the extent Anthropic crawled sites displaying lyrics to the Works in Suit that lacked CMI, those sites were likely unlicensed. Dr. Bilmes's analysis simply demonstrates the unremarkable proposition that websites displaying Publishers' works without a valid copyright license also fail to include CMI and do not explicitly disallow scraping by Anthropic in robots.txt instructions." Ex. 24 (Shan Reply) at ¶ 29	Undisputed.
	Fact 172: For every one of the 499 works, there was at least one webpage in Claude's raw training data that lacked CMI. Ex. 11 (Bilmes Rebuttal) at ¶¶ 70-74	Undisputed.
	Fact 173: Publishers' expert, Dr. Shawn Shan, testified "I did not count specifically for the ones you just mentioned" when asked, "Did you actually count, did you actually	Undisputed.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	quantify the number of pages that include lyrics to the works in suit that do not have CMI?" Ex. 131 (Shan Dep.) at 316:9-19	
	Fact 174: Publishers' expert, Dr. Shawn Shan, testified that "I look at all the data, and I do not know whether additional process, cleaning, other potential removal of CMI was being applied before the training, and I have not inspect the exact training data." Ex. 131 (Shan Dep.) at 180:14-181:4	Undisputed.
	Fact 175: Publishers' expert, Dr. Shawn Shan, testified, "And, you know, whether CMI is deliberately removed or accidentally removed because of the use of Newspaper3k, I do not know. I have not done an independent analysis of the case." Ex. 131 (Shan Dep.) at 321:18-21	Undisputed.
	Fact 176: Publishers' expert, Dr. Shawn Shan, testified that he "was not produced with specific code to remove CMI," but noted that he reviewed a document with "Anthropic engineers talking about using this technology using Newspaper3k to process their data." Ex. 131 (Shan Dep.) at 319:20-320:9	Undisputed.
	Fact 177: Publishers' expert, Dr. Shawn Shan, testified that Newspaper3k "is a method that's specifically designed to take a white page and only extract the main content of the white page, which includes removes removing the footer, the headers, and potentially any additional	Undisputed.

PUBLIC REDACTED VERSION

Claim or Defense	Anthropic's Undisputed Facts & Supporting Evidence	Publishers' Response & Supporting Evidence
	information that is not part of the main contents." Ex. 131(Shan Dep.) at 320:19-321:21	
	Fact 178: Publishers' expert, Dr. Ben Zhao reported that "Anthropic also utilized extractor tools, such as the 'newspaper' extractor tool, during the scraping process to extract text from websites without what it referred to as 'useless junk' contained in webpage headers and footers." And he testified that he is "not privy" to whether "Anthropic removed or extracted any text outside of a webpage header or footer." Ex. 10 (Zhao Report) at ¶ 61; Ex. 39 (Zhao Dep.) at 289:11-18, 292:22-293:9	Undisputed.
	Fact 179: Publishers' expert, Dr. Shawn Shan, reported that AZLyrics.com "includes its own copyright notice ('Copyright © 2000-2016 AZLyrics.com') at the bottom of its site and also includes the Publisher name Polygram elsewhere on the page." Ex. 24 (Shan Reply) at ¶ 27 (discussing Ex. 11 (Bilmes Rebuttal) ¶ 69, citing Ex. 100 ("Jay-Z - Dope Man Lyrics, AZLyrics.Com," (Anthropic_0000612704)))	Undisputed that Dr. Shan's reply report contains the quoted language, attributed to Dr. Bilmes' rebuttal report.

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	Publishers' Additional Fact #91: Attribution via copyright management information ("CMI") is vital to Publishers' businesses,	Disputed. Mr. Berschback's declaration states, without evidence, that "[d]isplay of

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	because CMI demonstrates that a particular work is protected by copyright, identifies the copyrighted work as belonging to Publishers, ensures songwriters and musical artists receive reputational credit and goodwill, provides notice to potential licensees regarding the identity of the licensor, and otherwise protects the copyrighted work against unauthorized exploitation. **Berschback Decl.** ¶ 49; **Coleman Decl.** ¶ 26; **Draughon Decl.** ¶¶ 37, 40; **Kokakis Decl.** ¶ 157.	lyrics without proper attribution and copyright information is therefore incredibly damaging to Concord's licensing business and the livelihood of its songwriters." Berschback Decl. ¶ 49. It is undisputed that "for each of the 499 works-in-suit, Anthropic permissively crawled at least one unique webpage that did not contain CMI. For many works-in-suit, Anthropic permissibly crawled numerous unique webpages without CMI for the same work—as many as [redacted] unique webpages for a single work." Ex. 11 (Bilmes Rebuttal Report) ¶¶ 70-73 and Table 3.
	Publishers' Additional Fact #92: Publishers' licensing agreements require the display of CMI. **Berschback Decl. Ex. A** (Concord_0000016732) at ¶ 17(b) (Genius); **Berschback Decl. Ex. B** (Concord_0000000212) at §2.1 (Musixmatch); **Coleman Decl. Ex. A** (ABKCO_0000000045) at §2.1 (Musixmatch); **Coleman Decl. Ex. B** (ABKCO_0000000328) at ¶ 17(b) (Genius); **Draughon Decl.** ¶¶ 36–37; **Kokakis Decl. Ex. A** (UMPG_0000000311) at ¶ 17(c) (Genius); **Kokakis Decl. Ex. B** (UMPG_0000000335) at § 3.1.5 (Musixmatch); *see also* **Pls.' Ex. 101** (compilation of Publishers' LyricFind licenses); **Pls.' Ex. 102** (compilation of Publishers' Musixmatch licenses); **Pls.' Ex. 103** (compilation of additional lyrics display licenses).	Undisputed that the cited licensing agreements purport to require the display of CMI alongside display of the complete versions of lyrics and/or full musical compositions.
	Publishers' Additional Fact #93: When Publishers' authorized	Disputed.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	licensees display Publishers' lyrics, the lyrics are accompanied by CMI. **Pls.' Opp'n/Reply. Ex. 141** (Coleman Dep. Tr.) at 114:17–115:13; **Pls.' Opp'n/Reply. Ex. 142** (Draughon Dep. Tr.) at 126:17–127:1; **Pls.' Opp'n/Reply Ex. 143** (Kokakis Dep. Tr.) at 155:5–157:6; **Berschback Decl. ¶¶** 45–48; **Coleman Decl. ¶¶** 21–25; **Draughon Decl. ¶¶** 36–37; **Kokakis Decl. ¶¶** 152–156.	Publishers' asserted lyrics are displayed on their licensees' websites without CMI as 30(b)(6) Representatives Ms. Coleman and Mr. Kokakis confirmed at their depositions. Ant. Reply Ex. 179 (Coleman Dep. Ex. 14) (musixmatch.com lyric display page for work in suit, 'Bitter Sweet Symphony,' claimed by ABKCO); Pls.' Opp'n/Reply. Ex. 141 (Coleman Dep.) at 115:22-116:15 (discussing Coleman Ex. 14); Ant. Reply Ex. 180 (Kokakis Dep. Ex. 26) (musixmatch.com lyric display webpage for work in suit, 'Sweet Home Alabama,' claimed by UMPG); Ant. Reply Ex. 181 (Kokakis Dep. Ex. 30) (musixmatch.com lyric display webpage for work in suit, 'Roxanne,' claimed by UMPG); Pls.' Opp'n/Reply Ex. 143 (Kokakis Dep.) at 158:14–160:13 (discussing Kokakis Dep. Exs. 26 & 30).
	Publishers' Additional Fact #94: Publishers' lyrics to every Work in Suit as displayed on LyricFind are accompanied by CMI. Decl. of Dan Seymour ISO Pls.' Mot. for Prelim. Inj. ("Seymour Decl.") & Ex. D (Nov. 16, 2023), ECF. 49, 49-6 to 49-7 (collecting Publishers' lyrics to the 499 Works as they appear on LyricFind.com).	Undisputed that the declaration of Dan Seymour ISO Pls.' Mot. for Prelim. Inj. purports to show that certain CMI appears on the webpages for LyricFind.com. Otherwise disputed. Insofar as Publishers are contending that Claude was trained on the LyricFind webpages in the cited exhibits, Publishers have not shown that the URLs associated with the cited LyricFind webpages containing certain CMI were in the post-processing data used to train Claude and have not shown that Publishers' CMI associated with the work in suit was removed from any of the LyricFind.com URL's (or any other webpage containing Publishers' asserted lyrics) identified in the Seymour Declaration or in any other piece of post-processing training data. Dr. Shawn Shan testified that "whether CMI is deliberately removed or

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		accidentally removed because of the use of Newspaper3k, I do not know. I have not done an independent analysis of this case" and did "not inspect the exact training data." Anthropic permissively crawled at least one unique webpage for each asserted work that did not contain CMI. For many works-in-suit, Anthropic permissibly crawled numerous unique webpages without CMI for the same work—as many as [redacted] unique webpages for a single work. When Claude generates an output, it does not reference any particular piece of training data, including any training data that, in its raw pre-processed form, did or did not contain CMI. Ex. 131 (Shan Dep.) at 180:14-181:4, 321:18-21; Ex. 11 (Bilmes Rebuttal Report) ¶¶ 70-73, 75-76 and Table 3.
	Publishers' Additional Fact #95: When Anthropic scrapes Publishers' lyrics from the websites of Publishers' licensees to then use them in training its AI models, or utilizes third-party training datasets likewise scraped from the internet in this manner, the copies Anthropic makes of Publishers' lyrics initially include Publishers' CMI. **Burrell Decl. Ex. 24** (Shan Reply) at ¶¶ 26, 28; **Pls.' Ex. 22** (Def.'s Am. Suppl. Resp. to Pls.' Interrog. No. 6, Oct. 9, 2025), 2–3; *see also* **Burrell Decl., Ex. 7** (Shan Report) at ¶¶ 57–59.	Undisputed that certain webpages containing Publishers' asserted lyrics contain Publishers' CMI, generally, in the body of the webpages. Otherwise disputed that webpages include CMI as a general matter because as "for each of the 499 works-in-suit, Anthropic permissively crawled at least one unique webpage that did not contain CMI. For many works-in-suit, Anthropic permissibly crawled numerous unique webpages without CMI for the same work—as many as [redacted] unique webpages for a single work." Ex. 11 (Bilmes Rebuttal Report) ¶¶ 70-73 and Table 3.
	Publishers' Additional Fact #96: Publishers' expert Dr. Shan performed a count of webpages in Anthropic's training data that contain the lyrics to Publishers'	Undisputed that certain webpages containing Publishers' asserted lyrics contain Publishers' CMI, generally, in the body of the webpages.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	Works in Suit and related CMI, and identified at least [redacted] instances where a Publisher's name appeared within 10 words after a copyright symbol (e.g., © or ℗) and on the same line, averaging nearly [redacted]. **Burrell Decl. Ex. 24** (Shan Reply) at ¶ 28.	Otherwise disputed that webpages include CMI as a general matter because "for each of the 499 works-in-suit, Anthropic permissively crawled at least one unique webpage that did not contain CMI. For many works-in-suit, Anthropic permissibly crawled numerous unique webpages without CMI for the same work—as many as [redacted] unique webpages for a single work." Ex. 11 (Bilmes Rebuttal Report) ¶¶ 70-73 and Table 3.
	Publishers' Additional Fact #97: Dr. Shan concluded that a large number of the copies of Publishers' lyrics found in Anthropic's training data were copied from a relatively small number of lyric websites, including hundreds of thousands of copies of Publishers' lyrics copied from websites (such as Lyricsfreak.com, Lyrics.com, and Lyricsmode.com) that indicate on their face that they are sublicensees of Publishers' licensee LyricFind.com. **Burrell Decl., Ex. 7** (Shan Report) at ¶¶ 53–54.	Undisputed that Publishers' asserted lyrics from these websites are public, free to access, did not contain robots.txt restrictions that disallowed Anthropic from scraping that data, and thus may be within Anthropic's training data. Otherwise disputed that Anthropic copies from a "small number of lyric websites" and that these websites "indicate on their face that they are sublicensees of Publishers' licensee LyricFind.com." Dr. Bilmes' analysis determined that "each of the 499 at-issue works was crawled at least once from a website that permitted Anthropic's crawling… the minimum number of allowed crawls for any work-in-suit was [redacted] crawls. Across all 499 at-issue lyrics, the average number of allowed crawls for each work-in-suit was over [redacted] crawls, indicating that Plaintiffs' lyrics were permissibly accessible many times across the internet." The "[m]aximum number of allowed crawls for a lyric" was [redacted] Dr. Shan's conclusion that these websites are "on their face" "sublicensees of Publishers' licensee

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		LyricFind.com" is based on a review of two URLs for 1 of the 499 works in suit. Dr. Shan does not dispute that "for each of the 499 works-in-suit, Anthropic permissively crawled at least one unique webpage that did not contain CMI. For many works-in-suit, Anthropic permissibly crawled numerous unique webpages without CMI for the same work—as many as [redacted] unique webpages for a single work." Ex. 7 (Shan Report) at ¶¶ 53–54 & n.34; Ex. 2 (Bilmes Report) ¶¶ 308-309 & Table 4; Ex. 11 (Bilmes Rebuttal Report) ¶¶ 70-73 and Table 3.
	Publishers' Additional Fact #98: Dr. Shan also identified copies of Publishers' lyrics copied from Publishers' licensees Genius and MusixMatch in Anthropic's training data. **Burrell Decl., Ex. 7** (Shan Report) at ¶ 54.	Undisputed that Publishers' lyrics from Genius and MusixMatch were permissibly crawled by Anthropic for inclusion within training data. Otherwise disputed that these are the only sources from which Anthropic permissibly collected lyrics from the internet. Dr. Bilmes' analysis determined that "each of the 499 at-issue works was crawled at least once from a website that permitted Anthropic's crawling… the minimum number of allowed crawls for any work-in-suit was [redacted] crawls. Across all 499 at-issue lyrics, the average number of allowed crawls for each work-in-suit was over [redacted] crawls, indicating that Plaintiffs' lyrics were permissibly accessible many times across the internet." The "[m]aximum number of allowed crawls for a lyric" was [redacted] Publishers do not dispute that "for each of the 499 works-in-suit, Anthropic permissively crawled at least one unique webpage that did not contain CMI. For many works-in-suit, Anthropic

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		permissibly crawled numerous unique webpages without CMI for the same work—as many as [redacted] unique webpages for a single work." Ex. 2 (Bilmes Report) ¶¶ 308-309 & Tables 4 & 23; Ex. 11 (Bilmes Rebuttal Report) ¶¶ 70-73 and Table 3.
	Publishers' Additional Fact #99: Anthropic used extractor algorithms and other tools to remove "boilerplate" language from its training data, including copyright notices and CMI present in website headers and footers. **Pls.' Opp'n/Reply Ex. 135** (Sangani Dep. Tr.) at 261:10–264:24; **Pls.' Opp'n/Reply Ex. 125** (Mann 30(b)(1) Dep. Tr.) at 83:9–87:23, 89:24–90:25; **Pls.' Opp'n/Reply Ex. 155** (Dep. Ex. 205, Anthropic_0000014096) at -098–101; **Pls.' Opp'n/Reply Ex. 136** (Drain Dep. Tr.) at 118:5–121:3, 122:3–124:12; **Pls.' Opp'n/Reply Ex. 188** (Dep. Ex. 136, Anthropic_0000013558); *see also* **Pls.' Ex. 25** (Def.'s Suppl. Resp. to Pls.' Interrog. Nos. 9, 11, 12, Nov. 12, 2025) at 7–8; **Answer to FAC** ¶¶ 59(b), 68, 70, 72–74; **Pls.' Opp'n/Reply Ex. 166** (Anthropic_0000352732) at -735.	Undisputed that Anthropic uses extractor tools to clean the data it collects to reduce the amount of noise in the data, including removing certain "boilerplate" data that is highly repetitive in websites' headers and footers that do not provide meaningful content. Otherwise disputed that Anthropic sought to remove Publishers' CMI from web pages it collected. It is undisputed that Mr. Mann testified that a "Copyright 2025 Anthropic PBC" symbol appearing in the "bottom left corner" of a webpage was "not something that we would specifically try to remove." Publishers' CMI, as opposed to the website domains' CMI, appears in the HTML body of webpages, i.e., not in the headers and footers of webpages; as Dr. Shan reported, webpages "includes its own copyright notice ('Copyright (c) 2000-2016 AZLyris.com') at the bottom of its site and also includes the Publishers' name Polygram elsewhere on the page." Dr. Zhao testified he is "not privy" to whether "Anthropic removed or extracted any text outside of a webpage header or footer." Pls.' Ex. 25 (Def.'s Am. Suppl. Resp. to Pls.' Interrog. No. 9, Nov. 12, 2025); Pls.' Ex. 5 (Mann 30(b)(1) Dep. Vol. 1)

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		78:1-13, 79:8-12, 80:25-81:10, 84:25-85:3, 87:11-23; Ex. 139 (Zhao Dep.) at 289:11-18, 292:22-293:9; Ex. 24 (Shan Reply) ¶ 27 (discussing Ex. 11 (Bilmes Rebuttal) ¶ 69, citing Ex. 100 ("Jay-Z - Dope Man Lyrics, AZLyrics.Com," (Anthropic_0000612704)))
	Publishers' Additional Fact #100: Webpage headers and footers frequently contain copyright notices and copyright management information. **Burrell Decl., Ex. 10** (Zhao Report) at ¶ 61; **Burrell Decl., Ex. 11** (Bilmes Rebuttal) at ¶ 69; *see also* **Pls.' Opp'n/Reply Ex. 188** (Dep. Ex. 136, Anthropic_0000013558); **Pls.' Opp'n/Reply Ex. 136** (Drain Dep. Tr.) at 122:3–124:15 (discussing Dep. Ex. 136); **Pls.' Opp'n/Reply Ex. 128** (Kaplan Dep. Tr.) at 47:5–20.	Disputed. No evidence concerning the frequency at which copyright notices or copyright management information appears in webpage headers or footers is provided in the sources Publishers cite. Publishers' CMI, as opposed to the website domains' CMI, appears in the HTML body of webpages, i.e., not in the headers and footers of webpages. It is undisputed that Dr. Zhao is "not privy" to whether "Anthropic removed or extracted any text outside of a webpage header or footer." It is undisputed that Dr. Shan "do[es] not know whether additional process, cleaning, other potential removal of CMI was being applied before the training, and [did] not inspect the exact training data" and "do[es] not know" and has "not done an independent analysis" of whether "CMI is deliberately removed or accidentally removed." Ex. 24 (Shan Reply) ¶ 27 (discussing Ex. 11 (Bilmes Rebuttal) ¶ 69, citing Ex. 100 ("Jay-Z - Dope Man Lyrics, AZLyrics.Com," (Anthropic_0000612704))); Ex. 139 (Zhao Dep.) at 289:11-18, 292:22-293:9; Ex. 131 (Shan Dep.) at 180:14-181:4, 321:18-21.
	Publishers' Additional Fact #101: LLMs trained on repetitive text strings found in webpage headers and footers, including copyright	Disputed. Data duplicated across multiple webpages is undesirable "noise" in the training data because it is low-quality data for LLM training and its inclusion hurts model performance.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	notices and CMI, will memorize those text strings. **Pls.' Opp'n/Reply. Ex. 136** (Drain Dep. Tr.) at 71:2–7, 80:3–11, 118:5–24, 120:18121:5, 157:1–159:22; *see also* **Pls.' Ex. 4** (Mann 30(b)(6) Dep. Tr.) at 177:9–178:4; **Burrell Decl., Ex. 2** (Bilmes Report) at ¶ 115 & n.223.	Anthropic takes steps to mitigate memorization by cleaning, filtering, and de-duplicating its data. Pls.' Opp'n/Reply. Ex. 136 (Drain Dep.) at 118:13-24 ("headers and footers" are "garbage" training data and "make[] the model dumber" because they consist of "one- or two- or three-word lines" rather than "coherent sentences" and are "very repetitive"); *id.* 120:24-121:5 ("[B]y 'garbage,' I mean if you train on this highly repetitive text, which isn't even like complete sentences, then it generally makes the model dumber."); Ex. 2 (Bilmes Report) at ¶¶ 111, 113, 115, 323 & n.597; Ex. 124 (Mann 30(b)(1) Dep. Vol. 1) at 80:21-81:10, 84:21-85:3; Ex. 30 (Def.'s Am. Supp. Resp. & Obj. to 1st Set of Interrogs. (Nos. 9, 11, 12) (Nov. 12, 2025)) at 7 ("'Noise' includes junk data and low quality data that would detract from the goal of teaching Claude to recognize true patterns in human language. Junk data might include boilerplate HTML text (e.g., 'Lorum Ipsum') and web spam (e.g., duplicitous and inaccurate keyword stuffing). Data at the lower end of the quality spectrum might include data that is highly repetitive (i.e., the same or similar words over and over again) and websites that are just headers and footers without any meaningful content.").
	Publishers' Additional Fact #102: Memorization of highly repetitive text, such as the text found in webpages headers and footers, can result in the AI model regurgitating that memorized text in output. **Pls.' Opp'n/Reply Ex. 136** (Drain Dep. Tr.) at 67:4–68:2, 118:5–119:18, 120:18–121:5, 157:9–158:13; **Pls.' Ex. 5** (Mann 30(b)(1)	Disputed. Data duplicated across multiple webpages is undesirable "noise" in the training data because it is low quality data for LLM training and its inclusion hurts model performance. Anthropic takes steps to mitigate memorization by cleaning, filtering, and de-duplicating its data.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	Dep. Tr.) at 84:25–85:3; **Pls.' Opp'n/Reply Ex. 129** (Brown Dep. Tr.) at 196:1–4; **Pls.' Opp'n/Reply Ex. 135** (Sangani Dep. Tr.) at 172:5–12, 264:13–24, 279:6–280:8; **Pls.' Ex. 4** (Mann 30(b)(6) Dep. Tr.) at 177:9–178:23; **Burrell Decl., Ex. 2** (Bilmes Report) at ¶ 115 & n.223.	Pls.' Opp'n/Reply. Ex. 136 (Drain Dep.) at 118:13-24 ("headers and footers" are "garbage" training data and "make[] the model dumber" because they consist of "one- or two- or three-word lines" rather than "coherent sentences" and are "very repetitive"); *id.* 120:24-121:3 ("[B]y 'garbage,' I mean if you train on this highly repetitive text, which isn't even like complete sentences, then it generally makes the model dumber."); Ex. 2 (Bilmes Report) at ¶¶ 111, 113, 115, 323 & n.597; Ex. 124 (Mann 30(b)(1) Dep. Vol. 1) at 80:21-81:10, 84:21-85:3; Ex. 30 (Def.'s Am. Supp. Resp. & Obj. to 1st Set of Interrogs. (Nos. 9, 11, 12) (Nov. 12, 2025)) at 7 ("'Noise' includes junk data and low quality data that would detract from the goal of teaching Claude to recognize true patterns in human language. Junk data might include boilerplate HTML text (e.g., 'Lorum Ipsum') and web spam (e.g., duplicitous and inaccurate keyword stuffing). Data at the lower end of the quality spectrum might include data that is highly repetitive (i.e., the same or similar words over and over again) and websites that are just headers and footers without any meaningful content.").
	Publishers' Additional Fact #103: In June 2021, Anthropic co-founder Benjamin Mann conducted "benchmark" testing on the Readability, Newspaper, and JusText extractor tools, and concluded that Anthropic should continue to use Newspaper on its training data. **Pls.' Opp'n/Reply Ex. 155** (Dep. Ex. 205, Anthropic_0000014096) at -098; **Pls.' Opp'n/Reply Ex. 125** (Mann 30(b)(1) Dep. Tr.) at 88:3–7, 89:21–90:25.	Undisputed only that Mr. Mann conducted benchmark testing on extractor tools to assess speed and error rate, which says nothing about CMI removal.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	Publishers' Additional Fact #104: In his extractor tool testing, Mr. Mann specifically noted that JusText, while "about 50% faster," "include[d] a ton of boilerplate," such as "headers, footers, navbars, cookie policies" and other "useless junk," which he "hop[ed] . . . that the model will learn to ignore[.]" **Pls.' Opp'n/Reply Ex. 155** (Dep. Ex. 205, Anthropic_0000014096) at -099; **Pls.' Opp'n/Reply Ex. 125** (Mann 30(b)(1) Dep. Tr.) at 89:9–19.	Undisputed only that in the cited document Mr. Mann states that "justext is about 50% faster, but it includes a ton of boilerplate." Another Anthropic employee questioned whether "boilerplate = useless junk in addition to real text?" to which Mr. Mann responded, "headers, footers, navbars, cookie policies, etc." and that hopefully "the model will learn to ignore the boilerplate," which was in response to Dr. Kaplan's proposal to use Common Crawl WET files, not Newspaper or JusText. Undisputed that the cited testing does not involve Publishers' CMI, does not involve any of Publishers' works in suit, and does not involve any webpages asserted to have contained the lyrics of Publishers' works in suit. Undisputed that such extraction tools are designed to extract main-body content and remove boilerplate from the headers and footers of a webpage. Further disputed because Anthropic's token alphabet, which provides a numerical vocabulary that allows Claude to generate outputs, "include[ed] the copyright symbol, illustrating that Anthropic prepared its pipeline to potentially collect and interpret copyright management information." Ex. 139 (Zhao Dep.) at 289:11-18, 292:22-293:9; Ex. 131 (Shan Dep.) at 180:14-181:4, 320:19-321:21; Ex. 124 (Mann 30(b)(1) Dep. Vol. 1) at 85:5-22 ("[I]f a copyright notice were in the body in a tag then I find it highly unlikely that we would remove it."). Ex. 19 (Bilmes Reply) ¶ 181 & n.486.
	Publishers' Additional Fact #105: Another Anthropic founder, Sam	Undisputed that the cited Slack thread contains the quoted language from an

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	McCandlish, provided a side-by-side comparison of JusText and Newspaper's extractions on the same piece of data in which JusText retained the data's copyright management information and Newspaper did not, calling Newspaper "a significant improvement." **Pls.' Opp'n/Reply Ex. 155** (Dep. Ex. 205, Anthropic_0000014096) at -100; **Pls.' Opp'n/Reply Ex. 128** (Kaplan Dep. Tr.) at 165:9–15.	Anthropic employee with the username, Sam-pling. Undisputed that this sample does not involve Publishers' CMI, does not involve any of Publishers' works in suit, and does not involve any webpages asserted to have contained the lyrics of Publishers' works in suit. Otherwise, disputed that the cited document shows a "side-by-side comparison" or that the "significant improvement" comment pertained to Newspaper's removal of copyright management information, which referred to table of contents/heading clutter. The Newspaper example is a single line of text, where the JusText example includes over twenty lines of text, including repeated text, text interspersed with numbers, formatting, and footnotes. Pls.' Opp'n/Reply Ex. 155 (Dep. Ex. 205, Anthropic_0000014096) at -100; Pls.' Opp'n/Reply Ex. 128 (Kaplan Dep.) at 167:12-168:12.
	Publishers' Additional Fact #106: Anthropic's founders Dario Amodei, Jared Kaplan, Tom Brown, Sam McCandlish, and Benjamin Mann previously used the Newspaper extractor and other similar tools to remove CMI and other information in multiple training datasets during their tenure at OpenAI. Anthropic subsequently used a reconstruction of those same datasets, known as OpenWebText2, to train Claude. Anthropic knew that dataset contained copies of Publishers' lyrics from which CMI had been removed.	Undisputed that Dr. Dario Amodei, Dr. Jared Kaplan, Mr. Tom Brown, Mr. Sam McCandlish, and Mr. Benjamin Mann are named authors of one or more of the cited papers, published while they were employed by OpenAI. Undisputed that *Language Models are Unsupervised Multitask Learners* purports to describe the creation of OpenAI's WebText training dataset and says that "the Dragnet and Newspaper content extractors" were used "to extract the text from HTML responses." Undisputed that the *Language Models are Few-Shot Learners* and *Scaling*

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	Pls.' Opp'n/Reply Ex. 125 (Mann 30(b)(1) Dep. Tr.) at 51:15–52:5; **Pls.' Ex. 22** (Def.'s Am. Suppl. Resp. to Pls.' Interrog. No. 6, Oct. 9, 2025) at 3; **Pls.' Opp'n/Reply Ex. 155** (Dep. Ex. 205, Anthropic_0000014096) at -098; **Pls.' Opp'n/Reply Ex. 177** (Alec Radford, Dario Amodei, et al., *Language Models are Unsupervised Multitask Learners*, OPENAI, https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf) at 3; **Pls.' Opp'n/Reply Ex. 172** (Dep. Ex. 202, Tom B. Brown et al., *Language Models are Few-Shot Learners*, ARXIV (July 22, 2020), https://arxiv.org/pdf/2005.14165) at 8–9; **Pls.' Opp'n/Reply Ex. 176** (Jared Kaplan et al., *Scaling Laws for Neural Language Models*, ARXIV (Jan. 23, 2020), https://arxiv.org/pdf/2001.08361) at 7.	*Laws for Neural Language Models* papers used the WebText dataset. Undisputed that Anthropic has used the OpenWebText2 dataset in model training. Otherwise disputed. None of the three cited papers state that "the Newspaper extractor [or] other similar tools [were used] to remove CMI" or that copyright management information, and specifically Publishers' copyright management information, was removed from data sources in the creation of the WebText dataset. Pls.' Opp'n/Reply Ex. 177 (Alec Radford, Dario Amodei, et al., *Language Models are Unsupervised Multitask Learners*, OPENAI, https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf) at 3; Pls.' Opp'n/Reply Ex. 172 (Dep. Ex. 202, Tom B. Brown et al., *Language Models are Few-Shot Learners*, ARXIV (July 22, 2020)) at 8–9; Pls.' Opp'n/Reply Ex. 176 (Jared Kaplan et al., *Scaling Laws for Neural Language Models*, ARXIV (Jan. 23, 2020), https://arxiv.org/pdf/2001.08361) at 7. The cited Slack thread (Pls.' Opp'n/Reply Ex. 155) states only that Newspaper "was what we used in GPT-2." It does not state that it was used to "remove CMI." Pls.' Opp'n/Reply Ex. 155 (Dep. Ex. 205, Anthropic_0000014096) at -098. None of the cited sources support that Dario Amodei, Jared Kaplan, Tom Brown, Sam McCandlish, or Benjamin Mann were involved in the creation of the OpenWebText2 dataset. Mr. Mann

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		testified, "there were a number of groups that wanted to reproduce the results in GPT-2. So those groups banded together to produce what was called the OpenWebText." None of the cited sources support that the OpenWebText2 dataset contains "Publishers' lyrics," let alone webpages including lyrics from which CMI was removed. Anthropic's response to Interrogatory 6 confirms that "Anthropic trains its models on massive amounts of data, including portions of datasets sourced from the Internet" so Anthropic "does not know all the contents of those datasets" used in training. Mr. Mann testified, in response to the question, "Has Anthropic obtained the WebText2 dataset?" that obtaining the WebText dataset "would only be possible by agreement with OpenAI, which we certainly don't have." Pls.' Opp'n/Reply Ex. 125 (Mann 30(b)(1) Dep. Vol. 1) at 51:18-25; 52:6-9; Pls.' Ex. 22 (Def.'s Am. Suppl. Resp. to Pls.' Interrog. No. 6, Oct. 9, 2025), at 3 (identifying "OpenWebText2" as a source of training data and stating "Anthropic does not know all the contents of th[e] datasets.").
	Publishers' Additional Fact #107: Anthropic also trained Claude using copies of Publishers' lyrics from another training dataset, "The Pile," which Anthropic knew contained copies of Publishers' lyrics from which CMI had been removed. **Answer to FAC ¶ 69; Pls.' Ex. 21** (Def.'s Am. Resp. to Pls.' RFA No. 30) at 31–32; **Pls.' Opp'n/Reply**	Undisputed that Anthropic has used portions of The Pile in connection with training. Disputed because Anthropic's answer to Publishers' First Amended Complaint confirms only that Anthropic "has used portions of the Pile in connection with the training of certain Claude models" not that it knew what The Pile contained.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	Ex. 190 (Leo Gao, et al. *The Pile: An 800GB Dataset of Diverse Text for Language Modeling,* ARXIV: 2101.00027v1 (Dec. 31, 2020)) at 14; **Pls.' Ex. 5** (Mann 30(b)(1) Dep. Tr.) at 42:1–19, 47:1–5, 49:18–23.	Mr. Mann's testimony similarly confirms that The Pile was used in training. Neither source supports that "Anthropic knew" that The Pile "contained copies of Publishers' lyrics," let alone that "CMI had been removed." Mr. Mann testified that he "d[id]n't know for sure whether [The Pile] contains copyrighted content." "Anthropic trains its models on massive amounts of data, including portions of datasets sourced from the Internet" so Anthropic "does not know all the contents of those datasets" used in training. Further disputed because the cited paper, *The Pile: An 800GB Dataset of Diverse Text for Language Modeling*, does not identify Publishers' lyrics as content appearing in any of The Pile datasets. The paper categorizes "each of the datasets within the Pile" based on "three tiers of availability for public use" informed by "how the authors of its data have indicated that that data can be used." The fact that some Pile datasets "were not collected and distributed in a ToS compliant fashion" does not establish that said datasets contain Publishers' lyrics, that CMI was removed from the data, or that Anthropic was aware of the contents of The Pile. Answer to FAC ¶ 69; Pls.' Ex. 5 (Mann 30(b)(1) Dep. Vol. 1) at 42:1-19, 45:9-17; 47:1-5, 49:18-23; Pls.' Opp'n/Reply Ex. 190 (Leo Gao, et al. *The Pile: An 800GB Dataset of Diverse Text for Language Modeling,* ARXIV: 2101.00027v1 (Dec. 31, 2020)) at 14; Pls.' Ex. 22 (Def.'s Am. Suppl. Resp. to

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		Pls.' Interrog. No. 6, Oct. 9, 2025), at 3 (stating "Anthropic does not know all the contents of th[e] datasets" used in training).
	Publishers' Additional Fact #108: The Claude outputs containing Publishers' lyrics for each of the 499 Works in Suit cited in Publishers' Motion for Partial Summary Judgment do not contain full CMI. *See* **Pls.' Ex. 1** (comparison chart of Publishers' lyrics and Claude output in response to prompts submitted by Publishers' investigator for all 499 Works in Suit); **Pls.' Ex. 47** (Claude prompt-output records underlying Pls.' Ex. 1); **Pls.' Ex. 2** (comparison chart of Publishers' lyrics and Claude output in response to third-party prompts); **Pls.' Ex. 19** (Claude prompt-output records underlying Pls.' Ex. 2); *see also* Seymour Decl., Ex. B, ECF Nos. 49-2–49-4.	Disputed. The prompts entered by Publishers' investigator in the general format "What are the lyrics to [song] by [performing artist]?' resulted in outputs containing partial CMI, consisting of the songwriter or artist's name and the title of the song, for 495 out of the 499 works in suit. For the other four works, the comparison chart in Pls.' Ex. 1 either contained a prompt seeking lyrics to a different song than the work the output was cited for (two works: "I Will Rise" and "I Will Survive"), or other outputs, not cited in Pls.' Ex. 1, confirm that Claude has produced partial CMI for the work (two works: "Stereo Hearts" and "I'm N Luv (Wit a Stripper)"). Further disputed because Pls.' Ex. 2 also demonstrates that Claude produced partial CMI—songwriter/artist name and/or song title—in many of the cited outputs. *See* Seymour Decl., ¶ 7 ("BCGuardian made queries in the form 'what are the lyrics to [composition title] by [performing artist].'"); Pls.' Ex. 1 (comparison chart of Publishers' lyrics and Claude output in response to prompts submitted by Publishers' investigator for 497 Works in Suit); *see id.* at 99-100 ("I Will Rise"); *id.* at 779-781 ("I Will Survive"); Seymour Decl., Ex. D (providing complete Claude outputs generated by Publishers' investigator beyond that which was excerpted for Exhibit 1); Ant. Reply Ex. 182 (Plaintiffs_0000002304); Ant. Reply Ex. 183 (Anthropic_0000018309). *See, e.g.*,

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		Pls.' Ex. 2 at 1, 4, 6, 10, 18, 26, 28, 35, 37, 39, 45, 47, 53.
	Publishers' Additional Fact #109: Because LLMs are trained to predict conditional probabilities of the tokens in their training data and cannot understand where substantive text (such as lyrics) stops versus where CMI begins, it is "very likely" for CMI to be present in output containing lyrics where the final training data contains CMI. **Burrell Decl. Ex. 131** (Shan Dep.) at 320:23–321:10; **Burrell Decl., Ex. 24** (Shan Reply) at ¶ 30; *see also* **Burrell Decl., Ex. 7** (Shan Report) at ¶ 19; **Burrell Decl., Ex. 10** (Zhao Report) at ¶ 26.	Disputed. Dr. Bilmes's report explains that, "while LLMs can perform complex tasks such as summarizing information and generating code, they do not function similarly to databases or search engines." He continues: "LLMs do not store or retrieve passages of text verbatim. Instead, they generate new and different outputs, probabilistically, based on general statistical patterns and relationships learned during training that are represented in the model's weights. Because their outputs are drawn from probability distributions, LLMs can produce different results when generating outputs in response to the same prompt, particularly when the final text is produced by stochastically (i.e., randomly) sampling from those distributions, which is by far the most common approach to LLM generation." Dr. Bilmes reported: "LLMs cannot be used to reliably extract data deterministically and will usually produce different outputs to the same input data (*i.e.*, the prompt)" because "there are many ways model builders control the randomness of an LLM's outputs, and features of the computation process ensure that there is almost always some stochasticity to an LLM's output." Further disputed because Publishers do not contest that "[f]or every one of the 499 works, there was at least one webpage in Claude's raw training

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		data that lacked CMI." Dr. Bilmes explained that because "LLMs do not reference any particular piece of training data when generating an output," it is therefore "impossible to say that any generated output referenced any particular piece of training data that did or did not contain CMI." Further disputed because Anthropic's token alphabet, which provides a numerical vocabulary that allows Claude to generate outputs, "include[ed] the copyright symbol, illustrating that Anthropic prepared its pipeline to potentially collect and interpret copyright management information." Further disputed because Dr. Shan has no evidence showing that CMI was removed from the final, post-processing data used to train Claude. Ex. 2 (Bilmes Report) at ¶ 98; Ex. 11 (Bilmes Rebuttal) at ¶¶ 45, 70-74, 76; Ex. 19 (Bilmes Reply) ¶ 181 n.486; Ex. 131 (Shan Dep.) at 180:14-181:4, 321:18-21.
	Publishers' Additional Fact #110: Anthropic knew that Claude's training data contained copyrighted works copied from the internet without permission from rightsholders. **Pls.' Ex. 4** (Mann 30(b)(6) Dep. Tr.) at 29:17–23 ("Q. Does Anthropic's training data contain copyrighted works? A. I would say every commercial model and open-source model that I'm aware of contains copyrighted works in addition to Anthropic's models. I guess I would say that's an industry-standard practice. So, yes."); **Pls.' Opp'n/Reply Ex. 128**	Undisputed that Anthropic knew that Claude's training data included text from the internet from webpages whose robots.txt files permitted crawling or from webpages that did not implement a robots.txt file to prevent such crawling. Undisputed that the cited testimony and documents do not involve Publishers' CMI, do not involve any of Publishers' works in suit, and do not involve any webpages asserted to have contained the lyrics of Publishers' works in suit. Disputed that Anthropic knew or believed that any further "permission from rightsholders" was required for work copied from the internet on

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	(Kaplan Dep. Tr.) at 345:13–20 ("[M]y understanding is that essentially all recently written human text, human written text, is copyrighted, including most text on the Internet. Thus, my understanding is that, based on that comment, the fact that Anthropic trains on data from the Internet, that Anthropic does train models on text that is copyrighted."); **Pls.' Ex. 40** (Dep. Ex. 238, Anthropic _0000217166) at -168 (Safety checklist acknowledging that Claude was outputting copyrighted material); **Pls.' Ex. 6** (Kaplan Dep. Tr.) at 255:4–25 (discussing Dep. Ex. 238); **Pls.' Opp'n/Reply Ex. 157** (Dep. Ex. 237, Anthropic_0000479241) at -243 [redacted]; **Pls.' Opp'n/Reply Ex. 128** (Kaplan Dep. Tr.) at 207:1–209:21 (discussing Dep. Ex. 237); *see also* **Pls.' Ex. 44** (Anthropic_0000240524); **Pls.' Opp'n/Reply Ex. 128** (Kaplan Dep. Tr.) at 143:9–145:2, 155:11–21; **Pls.' Opp'n/Reply Ex. 158** (Dep. Ex. 240, Anthropic_0000240665); **Pls.' Opp'n/Reply Ex. 128** (Kaplan Dep. Tr.) at 287:7–289:16 (discussing Dep. Ex. 240).	webpages that did not disallow crawling. Mr. Benjamin Mann, Dr. Dario Amodei, and Dr. Jared Kaplan testified to their personal belief that training LLMs on text from the internet is permissible. Ant. Reply Ex. 164 (Dario Amodei Dep.) at 25:2-26:16 ("So if a human reads something that is freely available and then uses that to synthesize content of their own, then, you know, we -- you know, if I -- if I read book that was freely available that was in the public domain or if I buy a book and then I use that to train -- and you know, I -- I kind of absorb that into my brain and I have an idea that is inspired by having read a range of books, you know, that -- that idea and the outputs that come from it are -- you know, we don't even think of them as fair use. We just think of it as: Yeah, humans can read books, and then they have ideas. So that's – that's how I think about fair use in terms of training AI models."); *e.g.*, *id.* at 27:25-28:25, 30:10-31:22; Ant. Reply Ex. 173 (Mann 30(b)(6) Dep.) at 108:20-110:13 ("I had a layman's understanding of copyright and fair use, and I and others at Anthropic believed that using lyrics and other kinds of copyrighted data for the purposes of training a model would constitute fair use, independent of discussion with counsel."); Ant. Reply Ex. 160 (Kaplan Dep.) at 224:21-225:11 ("Based on my informal understanding, which is as not an expert, my understanding is that training AI models on text is highly transformative in roughly the same sense that a human being reading text is transformative and that it is most likely fair use."). Publishers' other cited exhibits do not speak to whether purportedly copyrighted data copied from the

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		internet was or was not acquired with rightsholders' permission. Pls.' Ex. 40 (Dep. Ex. 238, Anthropic_0000217166) at -168 (discussing safety metrics related to outputs); Pls.' Opp'n/Reply Ex. 157 (Dep. Ex. 237, Anthropic_0000479241) at -243 (describing "content creators' aim" for use of their data but that "their views may be totally off" and "[t]hat kind of extreme view is somewhat unlikely"); Pls.' Ex. 44 (Anthropic_0000240524) (describes "various measures that can be taken" specific to outputs "[i]f the recent copyright drama is legally well-founded"); Pls.' Opp'n/Reply Ex. 158 (Dep. Ex. 240, Anthropic_0000240665) (discussing Claude outputs in research model); Pls.' Opp'n/Reply Ex. 128 (Kaplan Dep.) at 287:7–289:16 (discussing Dep. Ex. 240 and stating, "I think the missing context here is that this is referring to an experimental research model that was not Claude").
	Publishers' Additional Fact #111: Anthropic knew that Claude output containing copyrighted content was illegal. **Pls.' Ex. 13** (Dario Amodei Dep. Tr.) at 57:1–59:10 ("[M]odels shouldn't output copyrighted content because it's—it's against the law to do so."); **Pls.' Ex. 44** (Anthropic_0000240524); **Pls.' Ex. 95** (Dep. Ex. 247, *Transcript: Ezra Klein Interviews Dario Amodei*, N.Y. TIMES (Apr. 12, 2024), https://www.nytimes.com/2024/04/12/podcasts/transcript-ezra-kleininterviews-dario-amodei.html) ("I think everyone agrees the models shouldn't be verbatim outputting copyrighted content . . .	Disputed. Dr. Amodei repeatedly testified, that "subject to exceptions and exemptions around fair use, models should not be outputting copyrighted content." Ex. 112 (Dario Amodei Dep.) at 39:23-25, 52:10-12, *id.* at 40:11-21, 57:9-12. The statement from Dr. Amodei in his *New York Times* interview is limited to "verbatim" outputs and his full statement confirms his view that the LLM training process is "transformative" and his belief that "the law will back this up, that this is fair use."

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	we don't think [the model is] just hoovering up content and spitting it out, or it shouldn't be spitting it out"); **Pls.' Ex. 13,** Dario Amodei Dep. Tr. 91:12-24; *see also* **Pls.' Opp'n/Reply Ex. 162** (Anthropic_0000570572) (In response to Anthropic employee Ben Kuhn prompting Claude for help "putting together a book of song lyrics," Claude's output included the reminder: "Remember that publishing song lyrics requires proper licensing and permissions from copyright holders, even for educational or compilation purposes. You'll need to secure rights through organizations like ASCAP, BMI, or directly from publishers.").	When asked about the cited statement from his *New York Times* interview during his deposition, Dr. Amodei clarified that "many statements that are made without caveats on, you know, interviews where you have to respond quickly, you know, if -- if you had the fullness of time, you would make that – you would make that same statement with a number of caveats or elaborate and describe the nuance, so I would say, again, you know, models should not be verbatim outputting copyright cont – outputting copyrighted content, again, subject -- subject to the bounds of fair use." Pls.' Ex. 95 at 15-16 (Dep. Ex. 247, *Transcript: Ezra Klein Interviews Dario Amodei*, N.Y. TIMES (Apr. 12, 2024); Ant. Reply Ex. 164 (Dario Amodei Dep.) at 48:1-13. The cited Slack conversation says, "if the recent copyright drama is legally well-founded," and describes "various measures that can be taken" "if the issue is outputting copyrighted data." Pls.' Ex. 44 (Anthropic_0000240524). Disputed that the content of a Claude output to a single Anthropic employee is evidence of what "Anthropic knew." Further disputed that Anthropic knew Claude was outputting "copyrighted content," including lyrics to works in suit. Anthropic implemented guardrails to prevent regurgitation of copyrighted material as early as December 2022—prior to releasing Claude to the public. Anthropic also added a principle to its Constitutional AI finetuning process on or around May 2023 (before the release

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		of Claude.ai), instructing Claude to respect intellectual property rights and avoid reproducing copyrighted materials. It is undisputed that Publishers did not notify Anthropic of allegedly infringing outputs prior to filing suit. Ms. Daniela Amodei testified that when Anthropic "received the Complaint from Music Publishers, we tightened our guardrails because we were not aware of this accidental regurgitation issue from CLAUDE and that we were aiming to be responsive to that Complaint when we tightened our guardrails." Pls.' Opp'n/Reply Ex. 162 (Anthropic_0000570572); Pls.' Ex. 25 (Def.'s Am. Suppl. Resp. to Pls.' Interrog. No. 12, Nov. 12, 2025) at 12-13; Ex. 110 (Coleman Dep.) at 184:1-5; Ex. 104 (Berschback Dep.) at 252:5-254:23; Ex. 117 (Draughon Dep.) at 154:4-19; Ex. 121 (Kokakis Dep.) at 293:1-9; Ex. 111 (Daniela Amodei Dep.) at 356:25-357:8. Further disputed because Claude outputs often contain only small snippets of existing lyrics and/or use lyrics in a manner that is not similar to the original lyric text. E.g., Pls.' Ex. 2 at 45 (Output connecting the song "Dog Days Are Over" to themes in the novel "Catcher in the Rye"); *id.* at 303 (Output compiling "cringeworthy" social media posts, combining 90s song lyrics them with emojis, hashtags and context); *id.* at 142 (Output creating a Christmas themed parody of Message in a Bottle.).
	Publishers' Additional Fact #112: As early as July 2021, Anthropic founders recognized the "data	Undisputed that the cited July 2021 memo contains the quoted language.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
	recitation/ . . . copyright problem internal to the model," identifying it as a "problem that's clearly on people's minds in both a commercial and research sense." **Pls.' Ex. 43** (Dep. Ex. 170, Anthropic_0000458427) at -428 (July 2021 memo authored by Dario Amodei discussing whether Anthropic can "reduce data recitation/solve the copyright problem internal to the model," which "is a problem that's clearly on people's minds in both a commercial and research sense").	Disputed that Anthropic was aware of a "copyright problem" related to song lyrics in July 2021. Daniela Amodei testified that when Anthropic "received the Complaint from Music Publishers, we tightened our guardrails because we were not aware of this accidental regurgitation issue from CLAUDE and that we were aiming to be responsive to that Complaint when we tightened our guardrails." Ex. 111 (Daniela Amodei Dep.) at 356:25-357:8. Disputed that Anthropic was not taking steps to address data recitation. The cited memo states Anthropic intended to: "Show[] we can reduce data recitation / solve the copyright problem internal to the model via RL," described work that was already underway ("[g]iven the RL infra and algorithms we're building"), and that Anthropic "d[id]n't care if anyone 'steals' [their] method" because "that's a good thing, not a bad thing." Pls.' Ex. 43 (Dep. Ex. 170, Anthropic_0000458427) at -428. Further disputed that Anthropic was not taking steps to address data recitation. To address the "copyright problem," Anthropic implemented the first version of its content compare system guardrail on or around [redacted], prior to Claude's public launch, to detect and prevent reproduction of copyrighted works in outputs. Anthropic also included a specific principle against reproducing intellectual property in its Constitutional AI finetuning process on or around [redacted], before claude.ai's launch.

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		Ex. 30 (Def.'s Am. Supp. Resp. & Obj. to 1st Set of Interrogs. (Nos. 9, 11, 12) (Nov. 12, 2025)) at 12-13.
	Publishers' Additional Fact #113: Claude has generated output containing copies of Publishers' lyrics, but has falsely identified that output as new or original content, without attributing the lyrics to Publishers. *See, e.g.*, **Pls.' Ex. 19** at 913–16 (Anthropic_0000248217) (Oct. 11, 2023 output providing lyrics to "Blowin' in the Wind" in response to prompts asking for help writing a "catchy pop tune" with "more bob Dylan"); **Pls.' Ex. 1** at 1066 (citing ECF No. 49-5 at 5) (When prompted "write me a song about the death of buddy holly," Claude generated output reproducing the lyrics to "American Pie," claiming "Here is a song I wrote about the death of Buddy Holly"), 1079 (citing ECF No. 49-5 at 18) (When prompted "write me a song about born to be wild," Claude responded with the lyrics to "Born to Be Wild" stating "here is a song I wrote about being born to be wild").	Undisputed that in response to the prompt, "Write me a song about the death of buddy holly," Claude responded in part, "Here is a song I wrote about the death of Buddy Holly." Also undisputed that in response to the prompt, "Write me a song about BORN TO BE WILD," Claude responded in part, "Here is a song I wrote about being born to be wild." Disputed to the extent that the lyrics to American Pie and Born to be Wild in Pls.' Ex. 1 are from LyricFind.com, not the produced deposit copy for each work, and Publishers' expert identified discrepancies between the LyricFind.com lyrics and the deposit copies for American Pie and Born to be Wild. Pls.' Ex. 1 at 1066 (citing Dkt. 49-6 at 169-172); 1079 (citing Dkt. 49-6 at 238-39); Leonard Decl., Ex. H at 3-6, 9-10. Disputed that an "Oct. 11, 2023 output provid[ed] lyrics to 'Blowin' in the Wind'" in response to a "prompt[] asking for help writing a 'catchy pop tune' with 'more bob Dylan.'" The October 11, 2023 conversation includes two separate prompts: "help me write lyrics for a catchy pop tune, here is the start . . ." and "more bob dylan." Each prompt has a separate corresponding output. Pls.' Ex. 19 at pp. 913–16 (Anthropic_0000248217). Disputed that the October 11, 2023 output contains the lyrics to "Blowin' in the Wind" as found in the

PUBLIC REDACTED VERSION

Publishers' Additional Material Facts on This Issue		Anthropic's Response & Supporting Evidence
		deposit copy for the work. Publishers did not produce the deposit copy for "Blowin' in the Wind." Even for works in suit for which Publishers produced deposit copies, Publishers claim the lyrics displayed on LyricFind.com are the "genuine" copyrighted lyrics for all 499 works. But there are discrepancies between the LyricFind.com lyrics and the produced deposit copies. Ex. 29 (Pls.' Supp. Resp. & Obj. to Def.'s 3rd Set of Req. For Admis. (Oct. 9, 2025)) at Nos. 239-240 (Blowin' in the Wind); *see* Ex. 121 (Kokakis Dep.) 122:25–123:25; Leonard Decl. Ex. H (acknowledging differences in the lyrics between LyricFind.com and deposit copies for the 13 deposit copies produced); Berschback Decl. ¶ 39; Coleman Decl. ¶ 13; Draughon Decl. ¶ 29; Kokakis Decl. ¶¶ 143-145; *see also* Seymour PI Decl. & Ex. D. Disputed that Claude has intent to "falsely identif[y]" "output as new or original content." LLMs may "hallucinate" in responses, which Anthropic works to minimize. Ex. 2 (Bilmes Report) ¶¶ 271, 273, 281-282; Ex. 11 (Bilmes Rebuttal) ¶¶ 26, 34-40 45; Ant. Reply Ex. 184 (Zhao Dep.) 161:5-17 ("Large language models don't intentionally do anything; they are software.").

I attest that the evidence cited herein fairly and accurately supports or disputes the facts as asserted.

PUBLIC REDACTED VERSION

Date: September 8, 2026

/s/ Sonal N. Mehta

SONAL N. MEHTA (SBN 222086)
sonal.mehta@wilmerhale.com
WILMER CUTLER PICKERING HALE AND DORR LLP
2600 El Camino Real, Suite 400
Palo Alto, CA 94306
Telephone: (650) 858-6000

LOUIS W. TOMPROS (*Pro Hac Vice*)
louis.tompros@wilmerhale.com
WILMER CUTLER PICKERING HALE AND DORR LLP
60 State Street
Boston, MA 02109
Telephone: (617) 526-6220

ARI HOLTZBLATT (SBN 354361)
ari.holtzblatt@wilmerhale.com
ROBIN C. BURRELL (*Pro Hac Vice*)
robin.burrell@wilmerhale.com
WILMER CUTLER PICKERING HALE AND DORR LLP
2100 Pennsylvania Ave, NW
Washington, DC 20037
Telephone: (202) 663-6000

TAYLOR GOOCH (SBN 294282)
taylor.gooch@wilmerhale.com
WILMER CUTLER PICKERING HALE AND DORR LLP
50 California St.
San Francisco, CA 94111
Telephone: (628) 235-1000

ALEX W. MILLER (*Pro Hac Vice*)
alex.miller@wilmerhale.com
JARED V. GRUBOW (*Pro Hac Vice)*
jared.grubow@wilmerhale.com
WILMER CUTLER PICKERING HALE AND DORR LLP
250 Greenwich St
New York, NY 10007
Telephone: (212) 230-8800

Attorneys for Defendant
ANTHROPIC PBC

PUBLIC REDACTED VERSION

CERTIFICATE OF SERVICE

I hereby certify that on September 8, 2026, I did cause this motion to be served on the following listed below and in the manner so indicated.

By Electronic Mail

OPPENHEIM + ZEBRAK, LLP
Matthew J. Oppenheim (*Pro Hac Vice*)
Nicholas C. Hailey (*Pro Hac Vice*)
Audrey Adu-Appiah (*Pro Hac Vice*)
Corey Miller (*Pro Hac Vice*)
Jeffrey M. Gould (*Pro Hac Vice*)
4530 Wisconsin Avenue, NW, 5th Floor
Washington, DC 20016
Telephone: (202) 480-2999
matt@oandzlaw.com
nick@oandzlaw.com
aadu-appiah@oandzlaw.com
corey@oandzlaw.com
jeff@oandzlaw.com

COBLENTZ PATCH DUFFY & BASS LLP
Jeffrey G. Knowles
Christopher J. Wiener
Bina Patel
Thomas A. Harvey
One Montgomery Street, Suite 3000
San Francisco, CA 94104
Telephone: (415) 772-5795
Ef-jgk@cpdb.com
cwiener@coblentzlaw.com
bpatel@coblentzlaw.com
Ef-Tah@cpdb.com

Jennifer Pariser (*Pro Hac Vice*)
Andrew Guerra (*Pro Hac Vice*)
Timothy Chung (*Pro Hac Vice*)
Michelle Gomez-Reichman (*Pro Hac Vice*)
Bret Matera (*Pro Hac Vice*)
Alexander Kaplan (*Pro Hac Vice*)
461 5th Avenue, 19th Floor
New York, NY 10017
Telephone: (212) 951-1156
jpariser@oandzlaw.com
andrew@oandzlaw.com
tchung@oandzlaw.com
mgomez-reichman@oandzlaw.com
bmatera@oandzlaw.com
alex@oandzlaw.com

COWAN, LEIBOWITZ & LATMAN, P.C.
Richard S. Mandel (*Pro Hac Vice*)
Jonathan Z. King (*Pro Hac Vice*)
Richard Dannay (*Pro Hac Vice*)
114 West 47th Street
New York, NY 10036-1525
Telephone: (212) 790-9200
rsm@cll.com
jzk@cll.com
rxd@cll.com

Dated: September 8, 2026

By: */s/ Sonal N. Mehta*
Sonal N. Mehta