

[Counsel on signature page]

**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

**JOINT DISCOVERY DISPUTE
STATEMENT REGARDING
ANTHROPIC'S PRODUCTION OF
GUARDRAIL INFORMATION**

FILED UNDER SEAL

**HIGHLY CONFIDENTIAL
ATTORNEYS' EYES ONLY PORTIONS
HIGHLIGHTED IN YELLOW**

Judge Eumi K. Lee
Magistrate Judge Susan van Keulen

1 [REDACTED]
2 [REDACTED]
3 [REDACTED]
4 [REDACTED]
5 Using this same terminology, Publishers issued RFP 89 requesting “[d]ocuments sufficient
6 to show the contents of [REDACTED]
7 [REDACTED] (emphasis added). Publishers’
8 Third Set of RFPs, January 27, 2025. Anthropic made various objections to this Request but did
9 not dispute its relevance nor and agreed to produce for inspection documents “sufficient to show
10 the contents of [REDACTED].” Anthropic’s Amended Responses and
11 Objections to Plaintiffs’ Third Set of RFPs dated March 27, 2025.

12 Publishers inspected Anthropic’s training data and source code from September 9–11,
13 2025. Anthropic failed to make [REDACTED] available during this inspection. The Parties
14 scheduled a second inspection for October 2-3. In advance of that inspection, Publishers’ counsel
15 wrote to confirm that “[a]ll *databases of copyrighted content* that were used as reference sets in
16 [REDACTED]” would be available (emphasis added). Anthropic
17 responded that it “intends to make [these materials] available in the inspection environment.”
18 J. Grubow Sept. 30, 2025 email. However, again this material was not made available.

19 Immediately thereafter, the Parties conferred on this issue. Because the content of [REDACTED] is
20 large, Anthropic initially *offered to produce a list of the songs contained in [REDACTED]* rather than
21 provide the full dataset for inspection. Publishers declined that offer and explained they needed to
22 see the [REDACTED] as they appear in the dataset but did not need to inspect the entire
23 dataset beyond the lyrics file. Anthropic confirmed it would provide *the lyrics*. After the call,
24 Anthropic stated it would provide the full [REDACTED] dataset for inspection. *See* D. Patel Oct. 8, 2025
25 email (“Anthropic is willing to produce the entire hash corpus for review in the inspection
26 environment.”).

27 However, during Publishers’ third inspection on October 29, Anthropic made available
28 only a table containing hash values for the songs in the [REDACTED] database. A hash value is an encrypted

1 alphanumerical representation of particular content. Because two pieces of content with the same
 2 hash value will almost certainly be the same content, hash values can be used to determine when
 3 two pieces of content (such as a potential Claude output and copyrighted song lyrics) are exact
 4 copies of each other. But, crucially, because a hash value itself is just a string of letters and
 5 numbers, a hash value alone does not reveal anything about the content that produced it. As
 6 Anthropic well knows, the table of hash values Anthropic made available is useless for identifying
 7 the songs and lyrics it included in its [REDACTED] corpus—the whole point of RFP 89 and which Anthropic
 8 agreed to provide. Anthropic has still not identified the songs it included in its [REDACTED] database.

9 Publishers promptly raised this issue again with Anthropic’s counsel. While Publishers
 10 maintain they are entitled to the lyrics for [REDACTED], they offered to compromise
 11 by accepting production (rather than inspection) of the lyrics for just the 499 Works in Suit, along
 12 with all associated metadata, and by accepting Anthropic’s previous offer to produce a list of the
 13 other songs in [REDACTED]. Anthropic responded it complied with RFP 89 by making the hash values
 14 available for inspection and that it would not produce anything further unless Publishers stipulate
 15 that the materials sought “are HIGHLY CONFIDENTIAL under the Protective Order” and that
 16 these materials “shall be used ‘in connection with this case only for prosecuting, defending, or
 17 attempting to settle this litigation ... [and] shall not be used for any other purpose’ including to
 18 advance any future claims.” J. Grubow Nov. 6, 2025 email (quoting ECF No. 292). Publishers
 19 responded that they would continue to abide by the requirement of the Protective Order but would
 20 not further stipulate regarding use of discovery produced in this litigation.

21 **B. The lyrics Anthropic included in its guardrails are relevant and responsive.**

22 The songs Anthropic added to its guardrails after Publishers filed this case are clearly
 23 relevant and, until the instant dispute, Anthropic never contended otherwise. With respect to the
 24 lyrics for the 499 Works in Suit, the particular text Anthropic used to design its system to avoid
 25 [REDACTED]

26 [REDACTED] The identities of the remaining [REDACTED] dataset are also highly relevant.
 27 In an attempt to rebut Publishers’ willfulness allegations, Anthropic injected the issue of these
 28 [REDACTED] songs into the case by asserting that it deployed a new bespoke guardrail system containing

1 without any additional stipulation. Having so agreed, Anthropic cannot weeks later refuse to
2 produce those lyrics without an additional stipulation. Anthropic has not raised any issue regarding
3 the confidentiality of the information sought. As both the Court and Anthropic are well aware,
4 there is already a Protective Order in the case.

5 Nor can Anthropic withhold responsive and relevant discovery in this litigation solely on
6 the basis that the responsive information might be relevant to some future litigation. Revealing a
7 guilty conscience, Anthropic apparently believes Publishers may use the contents of the  
8 dataset in a future lawsuit against it and wishes to forestall Publishers' ability to do so ahead of
9 time. But "[p]rivacy of proprietary information, not immunity from suit, [i]s the legitimate purpose
10 of the protective order." *In re Dual-Deck Video Cassette Recorder Antitrust Litigation.*, 10 F.3d
11 693, 696 (9th Cir. 1993) (reversing finding of contempt for violation of protective order when
12 party filed second lawsuit without disclosing any confidential information). It goes without saying
13 that Publishers will, of course, protect the confidentiality of any information properly designated
14 under the Protective Order in this case. But Anthropic's concern is not confidentiality but
15 liability—and Anthropic cannot pervert the Protective Order into a liability shield.

16 Unable to explain why the discovery sought is unjustified, Anthropic resorts to erroneously
17 and inappropriately attacking Publishers' counsel. Anthropic has no basis to claim that Publishers'
18 counsel breached its obligations under the Protective Order in *Bartz* at any time. Notably, at no
19 point has Anthropic moved for sanctions or sought any other relief for any such alleged breach.
20 More significantly, Anthropic has never even suggested that Publishers' counsel has to date
21 breached its obligations under the Protective Order in this case and merely raised alarm that this
22 might happen. The Court should ignore Anthropic's unbecoming attempts to shield itself from
23 liability by casting baseless aspersions on Publishers' counsel's integrity. Should Publishers seek
24 to use any protected information in a later suit, Publishers can move for an exception to the
25 Protective Order, Anthropic can oppose that motion, and the Court can rule on the specific facts
26 presented—a remedy Anthropic's demanded stipulation would foreclose. Similarly, Anthropic
27 reserves its right to challenge Publishers' use of protected information in the future, which the
28 Court than consider in the context of the specific use at issue. What is clearly *not* proper is

Anthropic’s attempt to challenge Publishers’ hypothetical future use of information—which Publishers have not even received—based on Anthropic’s own fear that the information is incriminating.

II. Anthropic’s Position

This dispute is, indeed, simple. There are lyrics to [REDACTED] songs represented as hash values in the [REDACTED] (called the hash corpus), including those that represent lyrics to the 499 Works-in-Suit here. Anthropic made the entire hash corpus available to Publishers because that is what they asked for. When Publishers changed tack (at the very end of discovery) and said what they actually wanted was the raw text files, Anthropic offered to make the dataset containing the text for the lyrics related to the 499 Works-In Suit available for Publishers to inspect. Publishers refused that offer, insisting that Anthropic provide the identity of the *other* [REDACTED] songs, i.e., songs that are *not* at issue in this case. But the specific identity of those songs has no relevance to this case. Rather, the requested discovery appears to be a transparent attempt by Publishers and their counsel Oppenheim + Zembrak (“O+Z”) to obtain confidential materials from Anthropic in this litigation to amass evidence for a different, future suit. Publishers have already indicated their plans for a future suit, stating in the reply to their motion to amend that “denying the Motion would compel Publishers to file their new claims in a separate action.” ECF 426 at 2. But whatever future case they may wish to bring, the Protective Order entered by this Court in this case prohibits them from using any confidential materials produced by Anthropic in this case to inform those future claims.

In light of the patent irrelevance here of the identity of the [REDACTED] works, but in the interest of compromise, Anthropic asked O+Z to confirm that they would not use this list of works to bring separate litigation. Tellingly, Publishers repeatedly refused, and their portion of this dispute statement all but admits that improper purpose. What is more, this is not the first time O+Z has advanced the view that it may permissibly use confidential materials from one case in another: Only a couple months ago, O+Z defended their ability to use confidential material produced in *Bartz v. Anthropic* (where O+Z has appeared on behalf of a different group of publishers and where the protective order contains materially identical language) to litigate this case. O+Z’s pattern of disregarding this clear protective order language must end. Anthropic will be filing a motion that

seeks to eliminate their confusion about their obligations. Anthropic respectfully requests that the Court deny Publishers' request for the list of [REDACTED] songs as irrelevant or order that Publishers and O+Z not use the list to develop claims for future suits as a precondition to receiving the list.

A. Anthropic Produced The Materials Publishers Requested

Publishers' statement is rife with inaccuracies and mischaracterization of what Anthropic agreed to produce. Publishers consistently and repeatedly—through its discovery requests, written correspondence, and oral communication—requested the **hash corpus**, as Publishers' statement above (using the term “hash” 26 times and never identifying any specific request for the raw lyric files) confirms. Publishers' RFP 89 does not seek the raw lyric files [REDACTED], but rather “the contents of ... the hash corpus.” Anthropic has never taken the position that all hash values are relevant to this litigation, yet Anthropic provided Publishers with the hash corpus they asked for, and never agreed to provide more.¹ A hash corpus literally means a body of hash values. And these hash values represent [REDACTED]. Publishers now complain that hash values are an “encrypted alphanumerical representation” of content, but that is the definition of what they requested when they asked for the “hash corpus.” Only *after* concluding their inspections did Publishers (for the first time) ask for “the complete raw text data used to generate the hashes for the songs.” J. Pariser Email (Oct. 31, 2025). But that is not the information requested in the RFP Publishers served.

B. The Identity Of The [REDACTED] Songs Is Not Relevant

Publishers now request the specific song titles and artist names for the [REDACTED] works added to the [REDACTED], which is a guardrails system used to prevent the regurgitation of copyrighted materials that had been in place long before Publishers filed suit. The specific song titles and artist names for *other* works are plainly irrelevant to the 499 Works-in-Suit and are cumulative of discovery already produced. Publishers opaquely claim this information somehow bears on “the issues in this case” because Anthropic added “only [REDACTED] songs” to the [REDACTED] even though “[t]ens of millions of new songs are released every year.” But even that argument turns on the number and

¹ Publishers misleadingly alter an email (at 2) to suggest that Anthropic agreed to provide “All *databases of copyrighted content*” used in CCS. It did not. Anthropic's responded it “intend[ed] to make the materials *noted in the first two bullets* available in the inspection environment”—the first bullet of which was the hash corpus.

not the identity of the [REDACTED] songs. Anthropic has explained that *how* those [REDACTED] other works were selected is both irrelevant to Anthropic’s response to the allegations regarding the 499 Works-in-Suit and is also privileged and work product information. *Why* more than [REDACTED] works were not selected is also privileged, and, to the extent there was any nonprivileged reason, Publishers asked and received responses to their questions in their 30(b)(6) deposition of Dr. Vinay Rao. *See, e.g.,* Rao Tr. 92:11-93:6 [REDACTED] [REDACTED]”). Anthropic has produced the entire hash corpus, and code demonstrating how the [REDACTED] works are used in the guardrails and Publishers had ample opportunity in the inspection environment and through their pre- and post-suit investigations to test the efficacy of Anthropic’s guardrails; anything else is irrelevant, privileged, and cumulative. That Publishers continue to pursue this irrelevant information when they have confined themselves to the 499 Works-In-Suit underscores that their reason for seeking this confidential information is to improperly gin up new claims.

C. Neither Publishers Nor O+Z May Use The List To Develop Future Claims

Despite the lack of relevance to the claims in *this* suit, and to avoid burdening the Court, Anthropic proposed a path forward by which it would produce the song titles and artist names Publishers requested so long as Publishers would stipulate that they would only use that information *in this case*. Tellingly, Publishers refused. They defend in this dispute statement that “Publishers may use the contents of the [REDACTED] dataset in a future lawsuit against [Anthropic]” based on a disingenuous reading of *In re Dual-Deck Video Cassette Recorder Antitrust Litigation*, 10 F.3d 693 (9th Cir. 1993). That statement is deeply troubling and improper for two reasons.

First, Publishers’ request for the identity of non-works-in-suit is a fishing expedition to develop new claims. “Fishing expeditions to discover new claims, however, are not permitted.” *Bryant v. Mattel, Inc.*, 2007 WL 5432959, at *3 (C.D. Cal. Apr. 19, 2007); *see also Teradyne, Inc. v. Astronics Test Sys., Inc.*, 2022 WL 18397125, at *6 (C.D. Cal. Dec. 22, 2022) (denying discovery “seeking evidence of potential copyright infringement not currently at issue” because “the Court need not allow Plaintiff to engage in a fishing expedition for new causes of action”).

Second, the Protective Order prohibits exactly what Publishers apparently plan to do. But

that prohibition is apparently insufficient because, as described below, this is not the first time that O+Z has interpreted this protective order language to allow them to use confidential materials produced in one case to litigate another. That means that Publishers and/or O+Z may be violating the Protective Order **right now** by using Anthropic’s confidential information to investigate claims for a future suit. So, Anthropic reasonably requires additional assurances that Publishers and O+Z will not use the contents of the [REDACTED] dataset to generate or support claims in a future lawsuit. Publishers do not challenge Anthropic’s designation of these materials as highly confidential materials. *See* J. Pariser Email (Nov. 7, 2025). Under the Protective Order in this case, such confidential materials are “Protected Materials,” for which the Protective Order provides: “A Receiving Party may use Protected Material that is disclosed or produced ... in connection with this case only for prosecuting, defending, or attempting to settle **this** litigation ... [and] **shall not be used for any other purpose....**” ECF 292 at 11-12. This language is modeled on the Northern District of California’s Model Protective Order.

It is hard to think of a clearer way to say that Protected Material in this case cannot be used to prosecute, defend, or settle **other** litigation. Indeed, courts have consistently interpreted the model language as prohibiting the use of confidential information learned through discovery to advance new claims in other litigation. *See Silicon Genesis Corp. v. EV Grp. E. Thallner GmbH*, 2023 WL 6882749, at *2-3 (N.D. Cal. Oct. 18, 2023) (the Model Protective Order’s “plain language ... precludes the use outside this litigation of confidential material produced in this litigation,” and “unambiguously forb[ade]” “using the [producing party’s] confidential material produced in this litigation to launch a foreign action.”); *Del Campo v. Am. Corrective Counseling Servs., Inc.*, 2007 WL 1848660, at *5 (N.D. Cal. June 27, 2007) (“us[ing] [a] document” from the case “in connection with another case” violated same language); *LifeScan Scotland, Ltd. v. Shasta Techs., LLC*, 2013 WL 4604746, at *2 (N.D. Cal. Aug. 28, 2013) (using “AEO Material to identify defendants’ distributors and send them the cease and desist letters” violated same language); *On Command Video Corp. v. LodgeNet Ent. Corp.*, 976 F. Supp. 917, 922 (N.D. Cal. 1997) (“using [confidential] information to file a separate lawsuit ... violated the plain terms” of similar order). Any other interpretation would render this material provision meaningless.

1 *Dual-Deck* is not to the contrary. The *Dual-Deck* protective order was entirely different
2 from the model language and the language in this case’s Protective Order, as it precluded use of
3 “All information produced in discovery”—not just confidential information—a literal reading of
4 which “would be absurd.” *Id.*; see *Silicon Genesis*, 2023 WL 6882749, at *3 (distinguishing *Dual-*
5 *Deck* on this ground). Moreover, *Dual-Deck* involved whether a party committing “three technical
6 violations” of a protective order could be held in contempt, or if the use was in “substantial
7 compliance” with a reasonable reading of the protective order. 10 F.3d at 695. But the question
8 here is the correct reading of the Protective Order going forward, not whether O+Z can hide behind
9 a “substantial compliance” defense to justify past violations. In any event, any future use of the
10 [REDACTED] works to prepare another suit would not be a mere “technical violation.”

11 Even more troublingly, O+Z’s approach here is not a one-off event. After O+Z appeared
12 in the *Bartz* litigation, O+Z sent an email to Anthropic’s counsel in this case stating: “We’ve
13 become aware that Anthropic is using the pirated LibGen and PiLiMi torrented files as part of the
14 guardrails to which Anthropic stipulated in this litigation.” N. Hailey Email (Sept. 4, 2025). When
15 pressed, O+Z admitted that it had learned this information—which it was leveraging in *this* case—
16 from “the *Bartz* summary judgment briefing” that was “filed under seal in that case.” See N.
17 Hailey Ltr. (Sept. 12, 2025). O+Z defended doing so by misconstruing *Dual-Deck* as “holding
18 that the use of information gained by an attorney under a protective order in one case may
19 appropriately be used by the same attorney to develop a strategy applicable to a second action,”
20 just as they do here. *Id.* Publishers’ Counsel’s open admission that it interprets the language in
21 this Protective Order (and the near identical one in *Bartz*) to allow them to use confidential
22 materials to develop new claims is extremely concerning.

23 Given O+Z’s consistent and candid view that they interpret the Protective Order to allow
24 them to use confidential materials obtained in one case to develop claims for another, Anthropic
25 will be moving to confirm the meaning of the Protective Order, to put this broader practice to an
26 end. For this dispute, Anthropic respectfully requests that the Court either deny Publishers’ request
27 for the list of [REDACTED] songs as irrelevant or order that Publishers and O+Z not use any list Anthropic
28 is required to produce to develop claims for future suits.

Dated: November 19, 2025

By: /s/ Timothy Chung

OPPENHEIM + ZEBRAK, LLP

Matthew J. Oppenheim

Nicholas C. Hailey

Jeffrey M. Gould

Corey Miller

Audrey L. Adu-Appiah

(admitted pro hac vice)

4530 Wisconsin Ave., NW, 5th Floor

Washington, DC 20016

Telephone: (202) 480-2999

matt@oandzlaw.com

nick@oandzlaw.com

jeff@oandzlaw.com

corey@oandzlaw.com

aadu-appiah@oandzlaw.com

Alexander Kaplan

Jennifer Pariser

Andrew Guerra

Bret Matera

Timothy Chung

Michelle Gomez-Reichman

(admitted pro hac vice)

461 5th Avenue, 19th Floor

New York, NY 10017

Telephone: (212) 951-1156

alex@oandzlaw.com

jpariser@oandzlaw.com

andrew@oandzlaw.com

bmatera@oandzlaw.com

tchung@oandzlaw.com

mgomez-reichman@oandzlaw.com

COBLENTZ PATCH DUFFY & BASS LLP

Jeffrey G. Knowles (SBN 129754)

One Montgomery Street, Suite 3000

San Francisco, CA 94104

Telephone: (415) 391-4800

ef-jgk@cpdb.com

COWAN, LIEBOWITZ & LATMAN, P.C.

Richard S. Mandel

Jonathan Z. King

Richard Dannay

(admitted pro hac vice)

114 West 47th Street

New York, NY 10036-1525

Telephone: (212) 790-9200

rsm@cjl.com

jzk@cjl.com

rxk@cjl.com

Attorneys for Plaintiffs

Respectfully submitted,

By: /s/ Joseph R. Wetzel

LATHAM & WATKINS LLP

Joseph R. Wetzel (SBN 238008)

joe.wetzel@lw.com

Andrew M. Gass (SBN 259694)

andrew.gass@lw.com

Brittany N. Lovejoy (SBN 286813)

britt.lovejoy@lw.com

505 Montgomery Street, Suite 2000

San Francisco, California 94111

Telephone: +1.415.391.0600

Sarang V. Damle

(admitted pro hac vice)

sy.damle@lw.com

Sara Sampoli (SBN 344505)

sara.sampoli@lw.com

555 Eleventh Street NW, Suite 1000

Washington, DC 20004

Telephone: +1.202.637.2200

Allison L. Stillman

(admitted pro hac vice)

alli.stillman@lw.com

1271 Avenue of the Americas

New York, New York 10020

Telephone: +1.212.906.1747

Attorneys for Defendant

Respectfully submitted,

By: /s/ Sonal N. Mehta

**WILMER CUTLER PICKERING HALE AND DORR
LLP**

Sonal N. Mehta (SBN 222086)
sonal.mehta@wilmerhale.com
Allison Bingxue Que (SBN 324044)
allison.que@wilmerhale.com
2600 El Camino Real, Suite 400
Palo Alto, CA 94306
(650) 858-6000
Fax: (650) 858-6100

Ari Holtzblatt (SBN 354631)
ari.holtzblatt@wilmerhale.com
Robin C. Burrell (admitted *pro hac vice*)
robin.burrell@wilmerhale.com
2100 Pennsylvania Avenue NW
Washington, DC 20037
(202) 663-6964

Louis W. Tompros (admitted *pro hac vice*)
louis.tompros@wilmerhale.com
60 State Street
Boston, MA 02109
(617) 526-6886

Joseph Taylor Gooch (SBN 294282)
taylor.gooch@wilmerhale.com
50 California Street
Suite 3600
San Francisco, CA 94111
(628) 235-1002

Attorneys for Defendant

SIGNATURE ATTESTATION

Pursuant to Civil L.R. 5-1(i)(3), I hereby attest that concurrence in the filing of this document was obtained from all other signatories of this document. I declare under penalty of perjury that the foregoing is true and correct.

Dated: November 19, 2025

/s/ Timothy Chung

Timothy Chung