BOIES SCHILLER FLEXNER LLP
David Boies (*pro hac vice*)
333 Main Street
Armonk, NY 10504
(914) 749-8200
dboies@bsfllp.com

**LIEFF CABRASER HEIMANN &
BERNSTEIN, LLP**
Rachel Geman (*pro hac vice*)
250 Hudson Street, 8th Floor
New York, New York 10013-1413
(212) 355-9500
rgeman@lchb.com

**CAFFERTY CLOBES MERIWETHER &
SPRENGEL LLP**
Bryan L. Clobes (*pro hac vice*)
135 S. LaSalle Street, Suite 3210
Chicago, IL 60603
(312) 782-4880
bclobes@caffertyclobes.com

*Counsel for Individual and Representative
 Plaintiffs and the Proposed Class*

(additional counsel included below)

JOSEPH SAVERI LAW FIRM, LLP
Joseph R. Saveri (SBN 130064)
601 California Street, Suite 1505
San Francisco, California 94108
(415) 500-6800
jsaveri@saverilawfirm.com

DICELLO LEVITT LLP
Amy Keller (*pro hac vice*)
10 North Dearborn Street, Sixth Floor
Chicago, Illinois 60602
(312) 214-7900
akeller@dicellolevitt.com

Matthew Butterick (SBN 250953)
1920 Hillhurst Avenue, #406
Los Angeles, CA 90027
(323) 968-2632
mb@butericklaw.com

UNITED STATES DISTRICT COURT

NORTHERN DISTRICT OF CALIFORNIA

SAN FRANCISCO DIVISION

RICHARD KADREY, et al., *Individual and Representative Plaintiffs*, v. META PLATFORMS, INC., *Defendant.*	Case No. 3:23-cv-03417-VC **PLAINTIFFS' OPPOSITION TO META'S MOTION TO DISMISS THIRD AMENDED CONSOLIDATED COMPLAINT**

TABLE OF CONTENTS

TABLE OF AUTHORITIES

Page(s)

<u>**Cases**</u>

Statutes

Rules

Other Authorities

<u>**MEMORANDUM OF POINTS AND AUTHORITIES**</u>

In moving to dismiss in part Plaintiffs' Third Amended Consolidated Complaint, Dkt. 407 ("TACC"), Defendant Meta Platforms, Inc. ("Meta") characterizes Plaintiffs' new claims as an attempt to "distract from [the] core issue" of fair use. Dkt. 413 at 1 ("Mot."). But these new claims are predicated on facts that strike at the heart of Meta's "fair use" defense, while also supporting two additional causes of action. ***First***, the TACC squarely states a claim under Section 1202(b)(1) of the Digital Millennium Copyright Act ("DMCA") because Meta's intentional removal of copyright management information ("CMI") from the multitude of copyrighted books it pirated "enable[d]," "facilitate[d]," and "concealed" Meta's use of those pirated works in training its large language models ("LLMs"). TACC ¶ 105. ***Second***, Plaintiffs state a claim under California Penal Code § 502(c)(2) ("CDAFA") because Meta deliberately engaged in one of the largest data piracy campaigns *in history* to acquire text data for its LLM training datasets, torrenting and sharing dozens of terabytes of pirated data that altogether contain many millions of copyrighted works, including Plaintiffs' works. The Court thus should deny Meta's motion to dismiss and allow these claims—which impact the rights of millions of copyright holders—to proceed.

I.   PLAINTIFFS STATE A VALID DMCA CLAIM

A. Plaintiffs Allege the Requisite Harm for Standing.

Meta argues that Plaintiffs lack standing to bring their DMCA claim because the TACC purportedly does not allege any cognizable injury. Mot. at 6. But Plaintiffs' DMCA allegations include concrete harms with "close historical or common-law analogue[s]" in copyright infringement and unjust enrichment. *See TransUnion LLC v. Ramirez*, 594 U.S. 413, 424 (2021).

1. Plaintiffs plead a concrete injury.

For standing purposes, an injury is concrete if it has a "close historical or common-law analogue." *TransUnion*, 594 U.S. at 424. The analogue need not be an "exact duplicate," and in deciding concreteness, "[c]ourts must afford due respect to Congress's decision to impose a statutory prohibition or obligation on a defendant, and to grant a plaintiff a cause of action to sue over the defendant's violation of that statutory prohibition or obligation." *Id*. at 424-25. Here,

Plaintiffs allege that Meta violated 17 U.S.C. § 1202(b)(1) by removing CMI from its LLM training data *en masse* to enable and facilitate training on pirated works, including Plaintiffs' works, while simultaneously concealing the fact that Meta used these copyrighted works for training. TACC ¶ 107; *see also* ¶ 90 (removing CMI from training data "helps Meta conceal the copyrighted data on which Meta has trained, as the models cannot regurgitate data they are not trained on, and are circumscribed by their inputs"). Plaintiffs suffered an injury from Meta's CMI stripping in a manner analogous to at least two "harms traditionally recognized as providing a basis for lawsuits in American courts": copyright infringement and unjust enrichment. *See TransUnion*, 594 U.S. at 425.

a. Copyright infringement

The DMCA's prohibition of CMI removal from a copyright-protected work is analogous to copyright infringement—Plaintiffs' original claim in this case, and one that has been extensively developed throughout fact discovery. The Ninth Circuit credits Congress's judgment about which statutory violations constitute cognizable harm, especially when technological advances warrant legal protections against new harms. *See Van Patten v. Vertical Fitness Grp.*, 847 F.3d 1037, 1043 (9th Cir. 2017) (recognizing Congress's role in "elevating" injuries to concrete and cognizable status and holding valid the "substantive right to be free from" autodialed calls and texts under the TCPA); *Wakefield v. ViSalus,* 51 F.4th 1109, 1118 (9th Cir. 2022) (*Van Patten*'s analysis is consistent with *TransUnion*), *cert. denied*, 143 S. Ct. 1756 (2023). Congress enacted § 1202(b)(1)'s prohibition of CMI stripping for this exact reason: to give additional teeth to the Copyright Act's long-standing protections to copyright holders in light of the new risks posed by the Internet Age.[1] Indeed, both the Copyright Act and the DMCA protect similar rights involving copyright-protected works, and both grant the copyright owner the sole prerogative to decide how future iterations of the work may differ from the version the owner published. *See* 17 U.S.C. § 106

---

[1] *See* S. Rep. No. 105–190, at 8 (May 11, 1998) (DMCA addresses the "hesitat[ion]" of copyright owners "to make their works readily available on the Internet without reasonable assurance that they will be protected against massive piracy"); H.R. Rep. No. 105–551, pt. 1, at 9–10 (May 22, 1998) (DMCA "adapt[s]" law to "evolution in technology" that allows users "to send and retrieve perfect reproductions of copyrighted material easily and nearly instantaneously" worldwide).

(listing exclusive rights); 17 U.S.C. §§ 106(1), (2). Congress also provided similar remedies for DMCA violations as it did for copyright infringement under the Copyright Act: in both, the plaintiff can elect to recover for actual damages and profits, or alternatively, to seek statutory damages. *Compare* 17 U.S.C. §§ 504(b), (c) (copyright infringement) *with* 17 U.S.C. § 1203(c) (DMCA violations).

Given this close analogue, Plaintiffs allege a concrete injury: Meta's interference with their exclusive right to control their copyrighted works by removing the CMI from those works. That's the same right copyright holders enjoy, and it's the same injury copyright infringement creates. *See eBay Inc. v. MercExchange, LLC*, 547 U.S. 388, 392 (2006) ("a copyright holder possesses the right to exclude others from using his property") (cleaned up). On the facts pled here, Meta's DMCA violation (stripping Plaintiffs' CMI from its training data) occurred as part of the copyright infringement itself (repeatedly copying Plaintiffs' works during the LLM training process). Thus, copyright infringement serves as an appropriate "close historical analogue" of harm for a violation of § 1202(b)(1).

b. Unjust enrichment

Meta's § 1202(b)(1) violation (as alleged in the TACC) also causes injury analogous to unjust enrichment—a harm giving separate rise to Article III standing. The DMCA allows plaintiffs to disgorge "any profits of the violator that are attributable to the violation," and separate monetary damages aren't required. 17 U.S.C. § 1203(c)(2). Those harms are sufficiently analogous to common law harms conferring standing under Article III. "A person who is unjustly enriched at the expense of another is subject to liability." Restatement (Third) of Restitution and Unjust Enrichment § 1; *see also id.* cmt. a (enrichment "at the expense of another" requires "violation of the other's legally protected rights, without the need to show that the claimant has suffered a loss").

Congress recognized this analogy by including disgorgement relief in the DMCA that mirrors § 504(b) of the Copyright Act, and courts in turn recognize copyright's role in protecting against unjust enrichment. *See, e.g.*, *Nintendo of Am., Inc. v. Dragon Pac. Int'l*, 40 F.3d 1007, 1011 (9th Cir. 1994) (disgorgement under Copyright Act exists "to prevent the defendant's unjust

enrichment"); *Fair Isaac Corp. v. Fed. Ins. Co.*, 468 F. Supp. 3d 1110, 1116 (D. Minn. 2020) (holding "unjust enrichment" was among the purposes of § 504(b)'s disgorgement provision, and noting actual damages aren't necessary to award profits). Indeed, unjust enrichment has conferred standing to plaintiffs suing this exact defendant. *See In re Facebook, Inc. Internet Tracking Litig.*, 956 F.3d 589, 599-600 (9th Cir. 2020) ("California law recognizes a right to disgorgement of profits resulting from unjust enrichment, even where an individual has not suffered a corresponding loss. . . . Because California law recognizes a legal interest in unjustly earned profits, Plaintiffs have adequately pleaded an entitlement to Facebook's profits from users' personal data sufficient to confer Article III standing.").

Here, Meta's stripping of CMI to enable, facilitate, and conceal Meta's copyright infringement contributed to depriving Plaintiffs of payment for the copying and use of Plaintiffs' works. *See* TACC ¶ 109, Prayer for Relief ¶ (j); *cf. Raw Story Media v. OpenAI*, 2024 WL 4711729, at *5 (S.D.N.Y. Nov. 7, 2024) (plaintiffs "truly seek redress" for defendant's "*use* of [their] articles"). Further, Meta's CMI stripping saved on resources expended on the LLM training process because removal of CMI prevents its regurgitation—otherwise, Meta would be required to spend resources on separate post-training mitigations to prevent or limit this chronic regurgitation. *See* Restatement (Third) of Restitution and Unjust Enrichment, § 1 cmt. d (unjust enrichment applies in cases of "saved expenditure"). *See* TACC ¶ 90 (removal of CMI "clean[s]" datasets for "easier 'ingestion,'" among other purposes).

2. Courts have found DMCA standing under analogous facts.

Another court recently allowed a DMCA claim to proceed on similar AI training allegations. *See The Intercept Media v. OpenAI*, No. 24-cv-1515, Dkt. 122 (S.D.N.Y. Nov. 21, 2024) (Rakoff, J.). The allegations underpinning the DMCA claim in *Intercept Media* are similar in all relevant respects to those in this case, including with respect to harm. *Compare id.*, Dkt. 1, ¶¶ 51-61, 66-75 *with* TACC ¶¶ 80, 88-90, 103-109. The *Intercept Media* plaintiffs did not allege any harm or injury aside from copyright infringement itself, *see* No. 24-cv-1515, Dkt. 1, ¶¶ 59-61, 73-75, while also alleging that "[a]t least some of the time, ChatGPT provides or has provided

responses to users that regurgitate verbatim or nearly verbatim copyright-protected works of journalism without providing author, title, copyright, or terms of use information contained in those works," *id.* ¶ 35. In their motion to dismiss, the *Intercept Media* defendants argued plaintiffs' claims of injury were insufficient because they did not allege "whether ChatGPT has ever outputted" their works "to a user," *id.*, Dkt. 53 at 5-6, but Judge Rakoff nevertheless denied the motion and allowed plaintiffs' DMCA claim under § 1202(b)(1) to proceed. *See id.*, Dkt. 122.

Meta fails to address *Intercept Media*, instead focusing extensively on *Raw Story*, 2024 WL 4711729. But that case does not reach nearly as far as Meta contends.[2] To start, that district court predicated much of its decision on the fact that the plaintiffs failed to allege any "actual adverse effect stemming from this alleged DMCA violation." *Id.* at *4. In other words, those plaintiffs just alleged that CMI had been stripped, and claimed that as their injury. Here, however, Plaintiffs allege that Meta's removal of CMI eliminated the legal and business risk of training on copyrighted material such that the removal of CMI facilitated, enabled, and concealed Meta's own infringement—in other words, removal of CMI helped facilitate and conceal the very conduct underscoring Plaintiffs' Copyright Act claim itself. TACC ¶¶ 90, 107, 108. Further, the *Raw Story* plaintiffs *only* brought a § 1202(b)(1) claim, not a copyright infringement claim, and they raised no alternative theory of harm or unjust enrichment to the asserted § 1202(b)(1) violation. Finally, the *Raw Story* court allowed the plaintiffs to move to amend their DMCA claim to allege a cognizable injury, 2024 WL 4711729, at *5, and that motion is pending.

One last point: Meta analogizes to *Raw Story* based on the claim that "Plaintiffs still do not identify a single instance in which anyone has ever used Llama to *reproduce* Plaintiff's books without their CMI." Mot. at 7 (emphasis added). This argument should be ignored for two reasons. First, it is non-responsive to the TACC. Plaintiffs do not plead an "output" infringement claim in

---

[2] Additionally, Meta's reliance on *Doe 1 v. GitHub*, 672 F. Supp. 3d 837 (N.D. Cal. 2023) is misleading. The § 1202(b)(1) claim in that case was dismissed because, in Judge Tigar's view, the allegedly removed or altered information—unnamed "assertions, implications, and/or false descriptions of authorship or source"—*did not constitute CMI*, and the plaintiffs did not "plead facts that suggest such unidentified statements could constitute CMI." *Id.* at 859. Here, Meta does not dispute Plaintiffs' allegations that Meta systematically removed material that constitutes CMI.

---

this case, nor do they plead that Meta *distributed* works without CMI. Second, even if this red herring were relevant, *Raw Story* did not require dissemination as an absolute prerequisite for a § 1202(b)(1) claim, but only to the extent the plaintiffs in that case analogized their DMCA harm to property interference. 2024 WL 4711729, at *3. Regardless, to apply a dissemination requirement to § 1202(b)(1) is to ignore the plain language of the statute: Congress made "distribution" (*i.e.*, dissemination) a necessary element *only* for violations of subsections (2) and (3).[3] This Court should (and indeed, must) respect Congress's choice *not* to require distribution for a § 1202(b)(1) claim; Meta's unlawful CMI removal is sufficient. *See Van Patten*, 847 F.3d at 1043.

B. Plaintiffs State a Claim Under § 1202(b)(1).

To state a CMI stripping claim under § 1202(b)(1), Plaintiffs must plausibly allege that Meta (a) "intentionally remov[ed] or alter[ed]" CMI while (b) "knowing, or … having reasonable grounds to know, that it will induce, enable, facilitate, or conceal an infringement of any right under this title." 17 U.S.C. § 1202(b)(1). Here, Meta effectively concedes, as it must, that it intentionally stripped CMI from many millions of copyrighted works (Mot. at 10 n.5). Moreover, as the TACC alleges, the stripping of CMI was part of a scheme to *enable and facilitate* the use of copyrighted works as training data for Meta's LLMs, *i.e.*, to enable Meta's massive infringement of Plaintiffs' works, and to conceal evidence that it trained on these pirated works. Because the TACC also adequately alleges the requisite scienter, Plaintiffs sufficiently state a DMCA claim.

Meta's attack on Plaintiffs' DMCA claim is based on obfuscation and strawman arguments. **First**, Meta ignores Plaintiffs' allegations that its CMI removal was intended "to enable and facilitate infringement" (TACC ¶ 105), instead only addressing the alternative theory that removal was intended to "conceal an infringement." Strikingly, Meta even admits "that removal of CMI *facilitates* training" of its LLMs (Mot. at 8; *see also id*. at 1), *which is the very infringement Plaintiffs allege. See* TACC ¶¶ 88, 90 (stripping CMI made it "easier" for Meta to ingest and train

---

[3] Contrary to the reasoning in *Raw Story*, requiring dissemination for a § 1202(b)(1) removal claim would make § 1202(b)(1) redundant with § 1202(b)(3). Anyone who intentionally removed CMI from a copyrighted work in violation of § 1202(b)(1), and then disseminated that work, would necessarily have distributed it knowing that CMI had been removed, thus violating § 1202(b)(3).

---

its LLM on Plaintiffs' works); *Facilitate*, NEW OXFORD AMERICAN DICT. (3d ed. 2010) ("facilitate" is to "make (an action, process, etc.) easy or easier"). In other words, Meta recognizes that CMI removal as part of data processing helped to facilitate its use of millions of pirated works.

Similarly, Plaintiffs' allegations regarding concealment are sufficient to state a claim. Meta's contention that "the TACC does not cite to any documents or deposition testimony supporting Plaintiffs' theory that Meta removed CMI to aid or conceal infringement" relies on the false claim that the TACC was filed after "fact discovery has closed" on those issues. (Mot. at 8). Not so, as Meta well knows. Because Meta previously argued that CMI stripping was irrelevant, Judge Hixson had to order Meta to produce responsive documents on December 20, *after* the close of discovery. Dkt. 351 at 2-3 (ordering Meta to "produce documents and communications regarding Llama and stripping or removal of CMI from literary works"). On the substance, far from being "vague and nonsensical" (Mot. at 8), the allegations that Meta stripped CMI from its training data to conceal infringement are common sense. Because Meta's "models cannot regurgitate data they are not trained on" (TACC ¶ 90), Meta stripped CMI to ensure that Plaintiffs and others can't query its models to determine whether they were trained on certain copyrighted works (or reveal obvious evidence of the same, such as a © symbol). The TACC also alleges further steps Meta took to conceal its infringement after learning its Llama models "would generate incriminating answers to prompts inquiring about Llama's training." *Id.* ¶¶ 92-93.

Second, Meta's argument that Plaintiffs don't allege scienter with respect to concealment of infringement is wrong for the reasons set forth above: the TACC contains numerous "specific allegations" about Meta's knowing and intentional removal of CMI from Plaintiffs' works so Meta could more easily train its LLM on them. TACC ¶¶ 88-90, 105, 108. *See generally Baird v. ROK Drinks, LLC*, 2018 WL 5294867, at *2 (C.D. Cal. Jan. 9, 2018) (finding the DMCA does not have "nearly as strict an intent requirement as Defendants suggest," and all Plaintiff has to show is Defendant "[had] reasonable grounds to know" it would "conceal an infringement") (cleaned up).

Specifically, with respect to concealment, the TACC's allegations that Meta knew that "LibGen contained copyrighted books and articles," and that "removal of CMI would reduce the

chances that Plaintiffs and Class members would discover Meta had copied Plaintiffs' works or

used them to train Llama," TACC ¶ 108, are more than sufficient at the pleading stage, where

scienter need not be alleged with specificity: "Federal Rule of Civil Procedure 9(b) provides that

'intent, knowledge, and other conditions of a person's mind may be alleged generally.'" *Logan v.

Meta Platforms, Inc.*, 636 F. Supp. 3d 1052, 1064 (N.D. Cal. Oct. 25, 2022); *see APL Microscopic,

LLC v. Steenblock*, 2022 WL 4830687, at *2 (9th Cir. Oct. 3, 2022) (reversing dismissal of §

1202(b) claim where court could infer knowing removal of CMI). Courts in this circuit have

routinely denied similar motions to dismiss § 1202(b)(1) claims.[4] *See, e.g.*, *Baird*, 2018 WL

5294867, at *2-3; *Izmo, Inc. v. Roadster, Inc.*, 2019 WL 13210561, at *4 (N.D. Cal. Mar. 26, 2019)

("Mental conditions generally do not need to be alleged with specificity. Whether Roadster knew

or should have known that its activities would induce or enable . . . infringement . . . is more suited

to summary judgment."); *Daar v. Oakley, Inc.*, 2018 WL 9596129, at *3 (C.D. Cal. Sept. 27, 2018)

("*Stevens* was decided on summary judgment, not at the pleading stage. . . . A plaintiff's burden at

the pleading stage is not so exacting.") (cleaned up); *Beijing Meishe Network Tech. Co. v. TikTok,

Inc.*, 2024 WL 3522196, at *8 (N.D. Cal. July 23, 2024) (denying dismissal of a § 1202(b)(1) claim

and finding that the plaintiff sufficiently alleged the requisite mental state at the pleading stage).[5]

Faced with well-supported factual allegations of scienter, Meta is left to argue that its

"voluntary" disclosure of the Books3 pirated dataset for Llama 1 refutes Plaintiffs' allegations.

Mot. at 9-10. But Meta ignores its removal of CMI from tens of millions of copyrighted works

---

[4] *Mills v. Netflix, Inc.*, 2020 WL 548558 (C.D. Cal. Feb. 3, 2020) is inapposite because Netflix did not actually remove the plaintiff's name and Netflix attributed the plaintiff in the credits of the documentary. *Id.* at *3. Here, Meta has not provided any evidence to refute Plaintiffs' allegations that Meta removed CMI (because it cannot), nor did Meta ever attribute Plaintiffs for the use of their copyrighted works in training Llama. *Victor Elias Photography, LLC v. Ice Portal, Inc.*, 43 F.4th 1313 (11th Cir. 2022), is inapposite as an out-of-circuit case affirming summary judgment.

[5] Meta's reliance on two cases dismissing § 1202(b)(1) claims is misplaced. In *Andersen v. Stability AI Ltd.*, 700 F. Supp. 3d 853 (N.D. Cal. 2023), the court found plaintiffs' allegations conclusory because plaintiffs failed to identify (1) the type of CMI that was alleged to be removed and (2) which defendant stripped the CMI. *Id.* at 872. Neither problem is present here. And in *Philpot v. Alternet Media, Inc.*, 2018 WL 6267876 (N.D. Cal. Nov. 30, 2018), plaintiffs alleged only that "[defendant] should have known its alleged removal of the CMI … would induce, enable, facilitate, or conceal an infringement …" *Id.* at 5. Here, Plaintiffs provided abundant detail as to Meta's mental state behind removing the CMI of Plaintiffs' works. *See* TACC ¶¶ 88-94.

---

from other pirated datasets to train later Llama models: After Llama 1, Meta notably stopped identifying the datasets it used to train Llama "to avoid litigation from using copyrighted materials." TACC ¶¶ 79-80. Whether Plaintiffs were already aware of other infringing activity by Meta has no bearing on Meta's concerted scheme to enable, facilitate, and conceal evidence of *subsequent, ongoing, and different infringement*. Indeed, while Meta's expansive privilege assertions (subject to pending crime-fraud review, Dkt. 416) make it difficult to know for certain, a plausible inference is that after Meta's lawyers got involved, Meta sought to conceal its infringement both by hiding information about its datasets and stripping CMI. Meta's argument that there could be no concealment because "any datasets used (and any CMI removal) after this lawsuit was filed was subject to disclosure" warrants credit for boldness, though this Court should recall that Meta fought to shield such datasets from production; only some datasets were produced and only after Court intervention; and even once produced, they are still designated "AEO" such that even *Plaintiffs* (to say nothing of the public) cannot see them. *See* Dkts. 308 at 5, 351 at 2-3.

II.   PLAINTIFFS STATE A VALID CDAFA CLAIM

A. Plaintiffs' CDAFA Claim Is Not Preempted by the Copyright Act.

Plaintiffs' CDAFA claim is grounded in Meta's *illegal acquisition* of Plaintiffs' copyrighted works, while the Copyright Act claim is based on Meta's *illegal reproduction* of the works. *See generally hiQ Labs, Inc. v. LinkedIn Corp.*, 31 F.4th 1180, 1196 (9th Cir. 2022) ("[T]he CFAA [the federal analog to CDAFA] is best understood as an anti-intrusion statute and not as a misappropriation statute[.]") (quotations omitted). To be sure, the nature of Meta's *acquisition* is highly relevant to the copyright claim via bad faith, willfulness, and absence of fair use. But it is Meta's *reproduction* of the works that gives rise to the copyright claim itself. Plaintiffs' allegations that Meta used peer-to-peer file-sharing networks with pirated copyrighted works to acquire (and

share) massive quantities of those works, *see* TACC ¶¶ 87, 111,[6] are an "extra element" distinct

from the conduct—*i.e.*, the reproduction and use—underpinning Plaintiffs' copyright claim.

To determine whether state law claims are preempted by the Copyright Act, courts in this

Circuit must assess: (1) whether the subject matter of the state law claim falls within the subject

matter of copyright described in 17 U.S.C. §§ 102 and 103, and if so, (2) whether the rights asserted

under state law are equivalent to the rights contained in 17 U.S.C. § 106. *Laws v. Sony Music Ent.,*

Inc., 448 F.3d 1134, 1137-38 (9th Cir. 2006). "If a state law claim includes an extra element that

makes the right asserted qualitatively different from those protected under the Copyright Act, the

state law is not preempted by the Copyright Act." *Altera Corp. v. Clear Logic, Inc.*, 424 F.3d 1079,

1089 (9th Cir. 2005) (quotation marks omitted). A core purpose of copyright preemption "is to

prevent states from giving special protection to works of authorship that Congress has decided

should be in the public domain." *ProCD, Inc. v. Zeidenberg*, 86 F.3d 1447, 1453 (7th Cir. 1996).

In this case, the subject matter of the CDAFA claim does not "fall within subject matter of

copyright," and the rights asserted under CDAFA are not "equivalent to" those protected by the

Copyright Act.[7] Meta's unauthorized taking of Plaintiffs' data via participation in a worldwide

online piracy ring (as both downloader and distributor) is independently illegal. *See United States*

v. Christensen, 828 F.3d 763, 789 (9th Cir. 2015) ("A plain reading of [CDAFA] demonstrates

that its focus is on unauthorized taking or use of information."). CDAFA therefore has little to do

---

[6] A torrent is a "protocol . . . that is used to distribute a large computer file (such as of digitized music or video) that has been segmented in small pieces between a large number of peer-to-peer users." *Torrent*, MERRIAM-WEBSTER, https://www.merriam-webster.com/dictionary/torrent (last visited Feb. 10, 2025). IP pirates like Meta also upload or share files with others during (leeching) and after (seeding) downloading. *See, e.g.*, Izal et al., *Dissecting BitTorrent: Five Months in a Torrent's Lifetime*, INSTITUT EURECOM, § 2 (2004) ("BitTorrent distinguishes between two kinds of peers depending on their download status: clients that have already a complete copy of the file and continue to serve other peers are called *seeds*; clients that are still downloading the file are called *leechers*.") (emphasis in original).

[7] The elements of CDAFA and copyright infringement do not overlap. Even "access," which is required for both offenses, is defined differently in the copyright and CDAFA contexts. *Compare* CACI 1813 (CDAFA) *with* 9th Cir. Model Jury Instructions 17.18 (Copyright Infringement). Meta even admits that "CDAFA is primarily an 'anti-hacking statute intended to prohibit the unauthorized use of any computer system for improper or illegitimate purpose," Mot. at 10, while the rights conferred by the Copyright Act "assure contributors to the store of knowledge a fair return for their labors," *Harper & Row Pubs., Inc. v. Nation Enterprises*, 471 U.S. 539, 546 (1985).

---

with expanding copyright holders' rights in works that otherwise would be in the public domain—it instead prevents the unauthorized taking of data, including copyrighted works. Had Meta bought Plaintiffs' works in a bookstore or borrowed them from a library and then trained its LLMs on them without a license, it would have committed copyright infringement, but no CDAFA violation. Meta's decision to bypass lawful acquisition methods and become a knowing participant in an illegal peer-to-peer piracy network provides the "extra element" and is "qualitatively different" to establish an independent CDAFA violation. *See Capitol Audio Access, Inc. v. Umemoto*, 980 F. Supp. 2d 1154, 1160-61 (E.D. Cal. 2013) (rejecting preemption where plaintiffs alleged defendant "engage[d] in various activities relating to accessing computers and data without authorization").

Courts have sustained CDAFA and copyright claims based on the same course of conduct. *See Hidden Empire Holdings v. Angelone,* 2022 WL 17080131, at *5 (C.D. Cal. Sept. 30, 2022) (plaintiffs established likelihood of success on the merits of copyright and CDAFA claims); *DocMagic, Inc. v. Ellie Mae, Inc.,* 745 F. Supp. 2d 1119 (N.D. Cal. 2010) (holding plaintiff adequately pled both copyright infringement and CDAFA claims); *Ticketmaster LLC v. Prestige Ent. W., Inc.*, 315 F. Supp. 3d 1147 (C.D. Cal. 2018) (same). In *Hidden Empire*, the defendant logged into a film studio's computers without authorization to obtain an unreleased film (violating CDAFA) and then used the film to create a video game based on it (copyright infringement). 2022 WL 17080131, at *2. The CDAFA claim involved the method of acquiring the copyrighted work while the copyright claim involved copying the work. The same is true of Meta's peer-to-peer sharing, which made Plaintiffs' works available for upload to others, and Meta's independent copying of the works. The harm Plaintiffs suffered is the same type of harm courts have recognized as *not* preempted by copyright law. *See Capitol Audio Access*, 980 F. Supp. 2d at 1159-60 (acknowledging distinct copyright infringement and CDAFA harms from same course of conduct).

In contending otherwise, the sole case Meta cites for support lends no real support at all. As Meta conspicuously fails to mention, that court dismissed the CDAFA claim "*[b]ased solely on Plaintiffs' non-opposition*." *VBConversions, LLC v. Blueswitch, LLC*, No. 15-cv-09372, Dkt. 58 (C.D. Cal. Mar. 18, 2016) (emphasis added). Meta's attempt to mislead the Court in describing

---

VBConversions as "holding" that a CDAFA claim was preempted by the DMCA only serves to underscore the inadequacy of Meta's preemption argument. Meta cites no decision where a court actually held that the Copyright Act preempted a CDAFA claim, nor does any such decision exist.

Outside the CDAFA context, Ninth Circuit precedent similarly supports the conclusion that the unlawful acquisition of copyrighted content can give rise to various claims without raising the specter of preemption. *See G.S. Rasmussen & Associates v. Kalitta Flying Service*, 958 F.2d 896, 904, 906 (9th Cir. 1992) (conversion claim not preempted by Copyright Act where aircraft owner "pirated" engineer's aircraft modification system and then copied it to obtain FAA approval because "Federal copyright law governs only copying"); *Altera Corp.*, 424 F.3d at 1089-90 (holding "a state law tort claim concerning the unauthorized use of the software's end-product is not within the rights protected by the federal Copyright Act"); *Opperman v. Path, Inc.*, 84 F. Supp. 3d 962, 989 (N.D. Cal. 2015) (denying preemption where "Plaintiffs' allegations involve more than the mere reproduction of Plaintiffs' address book data; they include the unauthorized access, transmission, misuse, and misappropriation of that data."). The conversion analogue is instructive: Imagine a thief walks into Barnes & Noble and takes a book without paying for it. Then, he makes a copy of the stolen book. The theft clearly is an "extra element" distinct from the copyright infringement. The same is true here. Because the CDAFA claim includes the "extra element" of *acquisition,* the Copyright Act does not preempt it.

B. By Failing to Notify California's Attorney General, Meta Violated Rule 5.1's Notice Requirement.

Federal Rule of Civil Procedure 5.1 requires a party bringing a constitutional challenge to a state statute to *promptly* notify the state attorney general. Fed. R. Civ. P. 5.1(a). "The rule is designed to ensure that the entity with the greatest interest in protecting the constitutionality of the legislation – the sovereign that enacted it – has an opportunity to present all arguments it deems to be appropriate to defend its legislation." *Roe v. LexisNexis Risk Sols. Inc,* No. 2:12-cv-6284, Dkt. 23, at 5 (C.D. Cal. Jan. 3, 2013). As far as Plaintiffs can tell, Meta didn't provide such notice.

Because the preemption doctrine "has its roots in the Supremacy Clause," *Fid. Fed. Sav. & Loan Ass'n v. de la Cuesta,* 458 U.S. 141, 152 (1982), courts in this Circuit have held that preemption by a federal statute is a constitutional challenge for purposes of Rule 5.1's notice requirement. As one court explained, a preempted law is unconstitutional "at least in the circumstances where it is preempted." *See Douglas v. ReconTrust Co.,* 2012 WL 5470360, at *4 (W.D. Wash. Nov. 9, 2012); *see also ThermoLife Int'l LLC v. NeoGenis Labs Inc.,* 2020 WL 6395442, at *16 (D. Ariz. Nov. 2, 2020) ("Rule 5.1 should be interpreted as requiring notice" where party argued preemption by federal patent law). The State of California ("State") has a clear interest in ensuring CDAFA is interpreted correctly. The parties agree this dispute depends, at least in part, on which rights CDAFA protects. Accordingly, the State may choose to intervene to ensure the Court hears from the State on interpreting its own law. To hold that CDAFA is preempted by the Copyright Act without that opportunity would directly violate Rule 5.1(c). Accordingly, the Court may not enter a final judgment holding the statute preempted unless Meta provides notice to the State Attorney General, who will then have 60 days to intervene. *See* Fed. R. Civ. P. 5.1(c).

C. Plaintiffs State a Claim Under California Penal Code § 502(c)(2).

1. Plaintiffs have a possessory interest over the data Meta pirated.

Plaintiffs allege Meta torrented pirated datasets containing massive amounts of copyrighted works to train its LLMs—in other words, knowing copying of Plaintiffs' data without Plaintiffs' permission. Meta argues Plaintiffs haven't alleged sufficient harm or damages because they don't own the peer-to-peer file sharing network or data Meta accessed illegally. *See* Mot. at 13-15. But Meta's contention that Plaintiffs are not the "owners of data" under CDAFA is unsupported by any authority. Rather, an allegation that information accessed and copied by a defendant "carrie[s] financial value," as Plaintiffs allege here, supports CDAFA standing. *See, e.g.*, *Frasco v. Flo Health, Inc.*, 2024 WL 4280933, at *3 (N.D. Cal. Sept. 23, 2024) (denying summary judgment on CDAFA claim against Google for "obtaining and storing" plaintiffs' private health information from a third-party application); *Doe I v. Google LLC*, 2024 WL 3490744, at *1 (N.D. Cal. July 22, 2024) ("Google unlawfully tracks, collects, and monetizes [Plaintiffs'] private health information

through source code that is 'secretly embedded' on their health care providers' websites."); *see also Biden v. Ziegler*, 2024 WL 3498786, at *9 (C.D. Cal. June 20, 2024) (CDAFA does not "require[] a defendant to access a computer or device belonging to, or controlled by, Plaintiff").

Further, Meta's suggestion that its torrenting of copyrighted works by the millions can't be a CDAFA violation because the pirated datasets Meta accessed were "publicly available"—in addition to being preposterous—misreads the law and misapplies the facts. *See* Mot. at 14. The suggestion neglects that (1) CDAFA claims do *not* require circumventing technical barriers*, West v. Ronquillo-Morgan*, 526 F. Supp. 3d 737, 745–46 (C.D. Cal. 2020); (2) Meta proceeded with torrenting despite fully knowing doing so was illegal (TACC ¶¶ 87, 113); and, most importantly, (3) Meta's means of accessing the copyrighted works through torrenting means Meta also *participated in the peer-to-peer piracy networks* by making the pirated works available to other users worldwide (TACC ¶ 87). Addressing such wrongdoing falls within the identified purposes of CDAFA. Cal. Penal Code. § 502(c)(2) (creating liability for anyone who "[k]nowingly accesses and without permission takes, copies, or makes use of any data from a computer, computer system, or computer network").

2. Plaintiffs allege damage and loss resulting from Meta's rampant piracy.

Finally, Plaintiffs allege the exact type of damage or loss that sustains a CDAFA claim. Meta asserts that cognizable CDAFA damages must involve "the computer system, network, program, or data," and that Plaintiffs' alleged "lost royalties, reputational damages, and other consequential losses" are insufficient. Mot. at 15 (citations omitted). Yet in case after case on CDAFA, courts have held that any profit generated from a defendant's unauthorized access of plaintiffs' data confers a cognizable "damage or loss" to the plaintiffs. Here, Plaintiffs allege that they have a possessory interest in their copyrighted works, and those works can be readily monetized in the "vibrant market for content for AI training data—a market within which Meta participates." TACC ¶ 5. Further, Plaintiffs allege that by torrenting Plaintiffs' works from pirated databases in lieu of executing lawful licensing arrangements, Meta not only deprived Plaintiffs of *that* licensing revenue, but it also deprived Plaintiffs of *additional* revenue they could have

generated from "other users worldwide" because Meta simultaneously made the copyrighted works available to download by any interested internet user in the process of acquiring Plaintiffs' data. TACC ¶ 87. At the same time, Meta itself generated unjust profits through the massive cost savings associated with foregoing the need to license Plaintiffs' copyrighted works. ¶¶ 5, 111-113.

The Ninth Circuit has already held that this type of damage or loss is sufficient to support a CDAFA claim *against this exact defendant*: (1) Meta committed an "unauthorized use of [Plaintiffs'] information for profit"; (2) the information "carr[ies] financial value" through potential monetization opportunities; and (3) Plaintiffs "did not provide authorization for the use of their . . . information, nor did they have any control over its use." *In re Facebook*, 956 F.3d at 600-601. In other words, if Plaintiffs plausibly allege that there exists "a market for their data" and that Meta "misappropriated" that data through unauthorized computer access, those allegations are sufficient under CDAFA. *Brown v. Google LLC*, 685 F. Supp. 3d 909, 940 (N.D. Cal. 2023); *see also Rodriguez v. Google LLC*, 2025 WL 50425, at *10 (N.D. Cal. Jan. 7, 2025) (denying summary judgment to Google on CDAFA claim because "Plaintiffs have a stake in the value of their data" and "their data has economic value"). This type of loss has nothing to do with harm to computer infrastructure, but it has been repeatedly recognized in this District as sufficient for CDAFA claims, and Plaintiffs adequately pled it here.

CDAFA also authorizes additional categories of damages that bear no resemblance at all to Copyright Act damages. Just look at California's Civil Jury Instructions: CACI 1814 authorizes CDAFA compensatory damages for "money spent to investigate or verify whether [plaintiffs'] . . . data was or was not altered, damaged, or deleted by [defendant]." This investigative fee provision has no analogue in federal copyright law, and it constitutes a prime example of the "consequential losses" that Plaintiffs pled. TACC ¶ 114. Plaintiffs have engaged in protracted efforts to determine how Meta acquired their copyrighted works, where Meta shared that data, and how it used that data for itself. Those investigative costs are cognizable as CDAFA damages and are not preempted.

III. CONCLUSION

For these reasons, Plaintiffs request that the Court deny Meta's motion to dismiss.

Dated: February 11, 2025

By: /s/ *Maxwell V. Pritt*
          Maxwell V. Pritt

**LIEFF CABRASER HEIMANN &
BERNSTEIN, LLP**
Elizabeth J. Cabraser (SBN 083151)
Daniel M. Hutchinson (SBN 239458)
Reilly T. Stoler (SBN 310761)
275 Battery Street, 29th Floor
San Francisco, CA 94111-3339
(415) 956-1000
ecabraser@lchb.com
dhutchinson@lchb.com
rstoler@lchb.com

Rachel Geman (*pro hac vice*)
250 Hudson Street, 8th Floor
New York, New York 10013-1413
(212) 355-9500
rgeman@lchb.com

Kenneth S. Byrd (*pro hac vice*)
Betsy A. Sugar (*pro hac vice*)
222 2nd Avenue South, Suite 1640
Nashville, TN 37201
(615) 313-9000
kbyrd@lchb.com
bsugar@lchb.com

JOSEPH SAVERI LAW FIRM LLP
Joseph R. Saveri (SBN 130064)
Cadio Zirpoli (SBN 179108)
Christopher K.L. Young (SBN 318371)
Margaux Poueymirou (SBN 356000)
Holden Benon (SBN 325847)
Aaron Cera (SBN 351163)

601 California Street, Suite 1505
San Francisco, California 94108
(415) 500-6800

jsaveri@saverilawfirm.com
czirpoli@saverilawfirm.com
cyoung@saverilawfirm.com
mpoueymirou@saverilawfirm.com
hbenon@saverilawfirm.com
acera@saverilawfirm.com

BOIES SCHILLER FLEXNER LLP
David Boies (*pro hac vice*)
333 Main Street
Armonk, NY 10504
(914) 749-8200
dboies@bsfllp.com

Maxwell V. Pritt (SBN 253155)
Joshua M. Stein (SBN 298856)
44 Montgomery Street, 41st Floor
San Francisco, CA 94104
(415) 293-6800
mpritt@bsfllp.com
jstein@bsfllp.com

Jesse Panuccio (*pro hac vice*)
Jay Schuffenhauer (*pro hac vice*)
1401 New York Ave, NW
Washington, DC 20005
(202) 237-2727
jpanuccio@bsfllp.com
jschuffenhauer@bsfllp.com

Joshua I. Schiller (SBN 330653)
David L. Simons (*pro hac vice*)
55 Hudson Yards, 20th Floor
New York, NY 10001
(914) 749-8200
jischiller@bsfllp.com
dsimons@bsfllp.com

*Interim Lead Counsel for Individual and
Representative Plaintiffs and the Proposed Class*

Matthew Butterick (SBN 250953)
1920 Hillhurst Avenue, #406
Los Angeles, CA 90027
(323) 968-2632
mb@buttericklaw.com

**CAFFERTY CLOBES MERIWETHER &
SPRENGEL LLP**
Bryan L. Clobes (*pro hac vice*)
Alexander J. Sweatman (*pro hac vice*)
Mohammed A. Rathur (*pro hac vice*)
135 S. LaSalle Street, Suite 3210
Chicago, IL 60603
(312) 782-4880
bclobes@caffertyclobes.com
asweatman@caffertyclobes.com
mrathur@caffertyclobes.com

DICELLO LEVITT LLP
Amy Keller (*pro hac vice*)
Nada Djordjevic (*pro hac vice*)
James Ulwick (*pro hac vice*)
10 North Dearborn Street, Sixth Floor
Chicago, Illinois 60602
(312) 214-7900
akeller@dicellolevitt.com
ndjordjevic@dicellolevitt.com
julwick@dicellolevitt.com

David Straite (*pro hac vice*)
485 Lexington Avenue, Suite 1001
New York, New York 10017
(646) 933-1000
dsraite@dicellolevitt.com

**COWAN DEBAETS ABRAMS &
SHEPPARD LLP**
Scott J. Sholder (*pro hac vice*)
CeCe M. Cole
60 Broad Street, 30th Floor
New York, NY 10004
(212) 974-7474
ssholder@cdas.com
ccole@cdas.com

*Counsel for Individual and Representative
Plaintiffs and the Proposed Class*