

FOR PUBLICATION

**UNITED STATES COURT OF APPEALS
FOR THE NINTH CIRCUIT**

JOHN DOE 1; JOHN DOE 2,

Plaintiffs - Appellants,

v.

TWITTER, INC.; X CORP.,

Defendants - Appellees.

No. 24-177

D.C. No.
3:21-cv-00485-
JCS

OPINION

Appeal from the United States District Court
for the Northern District of California
Joseph C. Spero, Magistrate Judge, Presiding

Argued and Submitted February 3, 2025
San Francisco, California

Filed August 1, 2025

Before: M. Margaret McKeown, Danielle J. Forrest, and
Gabriel P. Sanchez, Circuit Judges.

Opinion by Judge Forrest

SUMMARY*

Communications Decency Act

The panel affirmed in part and reversed in part the district court's judgment, and remanded, in a case in which the district court dismissed a complaint in which Plaintiffs—two minor boys—sued Twitter after it slow-walked its response to reports about, and did not immediately remove from the platform, pornographic content that a trafficker had coerced Plaintiffs into producing.

Plaintiffs advanced numerous claims, all of which the district court dismissed, primarily based on immunity provided under § 230 of the Communications Decency Act of 1996, which protects internet-based publishers of third-party content from liability.

The panel held that Twitter is immune from liability on Plaintiffs' claim, under the Trafficking Victims Protection Reauthorization Act of 2003 (TVPRA), alleging that Twitter knowingly benefitted from a sex-trafficking venture in violation of 18 U.S.C. § 1591, giving rise to civil liability under 18 U.S.C. § 1595. The panel explained that any activity that can be boiled down to deciding whether to exclude material that third parties seek to post online is perforce immune under § 230, absent the exception set forth in the Allow States and Victims to Fight Online Sex Trafficking Act (FOSTA). FOSTA withdraws immunity for any claim in a civil action brought under § 1595 if the underlying conduct constitutes a violation of § 1591. The

* This summary constitutes no part of the opinion of the court. It has been prepared by court staff for the convenience of the reader.

panel held that FOSTA's exception to § 230 immunity does not apply because Plaintiffs did not allege that Twitter itself violated § 1591.

The panel held that Twitter is also immune from liability on Plaintiffs' California product-defect claim based on Twitter's failure to remove posts under review as being child pornography and its creation of search features that amplify child-pornography posts.

The panel therefore affirmed the district court's dismissal of the TVPRA claim and the California products liability claim as it relates to Plaintiffs' removal and amplification design-defect theories.

However, the panel held that Plaintiffs' claims for negligence per se and their product-liability theory based on defective reporting-infrastructure design are not barred by § 230 immunity because they do not arise from Twitter's role as a publisher. The panel therefore reversed the dismissal of these claims and remanded for further proceedings.

COUNSEL

Peter A. Gentala (argued), Benjamin W. Bull, Danielle B. Pinter, and Christen M. Price, National Center on Sexual Exploitation, Washington, D.C.; Lisa D. Haba and Adam A. Haba, The Haba Law Firm PA, Longwood, Florida; Paul A. Matiasic, The Matiasic Firm PC, San Francisco, California; for Plaintiffs-Appellants.

Derek L. Shaffer (argued), Quinn Emanuel Urquhart & Sullivan LLP, Washington, D.C.; Dylan Bonfigli, Quinn Emanuel Urquhart & Sullivan LLP, Los Angeles, California; for Defendant-Appellees.

S. Mary Liu, Aylstock Witkin Kreis & Overholtz, Alameda, California; Hillary Nappi, Hach Rose Schirripa & Cheverie LLP, New York, New York; Marci A. Hamilton and Jessica Schidlow, CHILD USA, Philadelphia, Pennsylvania; for Amici Curiae CHILD USA, Hillary Nappi Esq., and Mary Liu Esq..

OPINION

FORREST, Circuit Judge:

We are once again confronted with the scope of an interactive computer service provider’s liability for hosting illegal content—namely, child pornography. A trafficker coerced Plaintiffs—two minor boys—into producing pornographic¹ content that the trafficker then posted on Twitter. Plaintiffs sued Twitter after it slow-walked its response to reports about this content and did not immediately take it off the platform. In their 13-count complaint, Plaintiffs advanced numerous claims, all of which the district court dismissed, primarily based on immunity provided under § 230 of the Communications Decency Act of 1996.

Though expansive, there is nuance to § 230 immunity. Here, we conclude that Twitter is immune from liability on Plaintiffs’ claim that it knowingly benefitted from sex trafficking and on their product-defect claim based on Twitter’s failure to remove posts under review as being child pornography and its creation of search features that amplify child-pornography posts. These claims hinge on Twitter’s role as a publisher of third-party content, which triggers § 230. But Plaintiffs’ claims for negligence per se and their product-liability theory based on defective reporting-infrastructure design are not barred by § 230 immunity

¹ Plaintiffs refer to “child pornography” as “child sexual abuse material” to “better capture[] the harmful nature of the material.” Their point is well taken. Nonetheless, we employ the language of the relevant statutes, which refer to “child pornography.” *See, e.g.*, 18 U.S.C. § 2252A.

because they do not arise from Twitter’s role as a publisher. Thus, we affirm in part, reverse in part, and remand.

BACKGROUND

I. Section 230 Immunity

Congress enacted § 230 to protect internet-based publishers of third-party content from liability. *See* 47 U.S.C. § 230(a)–(b) (enumerating Congress’s objectives); *Fair Hous. Council of San Fernando Valley v. Roommates.com, LLC*, 521 F.3d 1157, 1162–63 (9th Cir. 2008) (en banc) (recounting the legislative backdrop). The law reads: “No provider . . . of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.” 47 U.S.C. § 230(c)(1). Based on this text, § 230 immunity protects only: “(1) a provider or user of an interactive computer service (2) whom a plaintiff seeks to treat . . . as a publisher or speaker (3) of information provided by another information content provider.” *Barnes v. Yahoo!, Inc.*, 570 F.3d 1096, 1100–01 (9th Cir. 2009).

A “publisher” is someone who “reviews material submitted for publication, perhaps edits it for style or technical fluency, and then decides whether to publish it.” *Id.* at 1102. A claim that “obliges the defendant to ‘monitor third-party content’” to avoid liability also treats the defendant as a publisher. *Calise v. Meta Platforms, Inc.*, 103 F.4th 732, 742 (9th Cir. 2024) (citing *HomeAway.com, Inc. v. City of Santa Monica*, 918 F.3d 676, 682 (9th Cir. 2019)).

These functions contrast with those undertaken by an “information content provider,” who “is responsible, in whole or in part, for the creation or development of information provided.” 47 U.S.C. § 230(f)(3). Thus, if an

interactive computer service provider is disseminating content that it created, it is functioning as a “content provider,” not a publisher, and has no immunity under § 230. *Roommates.com*, 521 F.3d at 1162–63 (citing 47 U.S.C. § 230(f)(3)). But if the provider is disseminating content created by others, it is functioning as a publisher and is immune from liability related to that content. *Id.* Of course, an interactive computer service provider can function as both a publisher and a content provider. Therefore, courts must carefully consider which status any given claim derives from. *Barnes*, 570 F.3d at 1102.

II. Plaintiffs’ Allegations

Because this appeal arises from an order granting Twitter’s motion to dismiss, we accept the Plaintiffs’ allegations as true. *See id.* at 1098 n.1.

When Plaintiff John Doe #1 was 13 years old, he communicated on Snapchat with an individual who Doe #1 thought was his classmate. The Snapchat user requested nude images from Doe #1 and his friend, Plaintiff John Doe #2, who both complied. In reality, the Snapchat user was a child-pornography trafficker who used the images to blackmail Plaintiffs into producing additional pornographic images.

Eventually, Plaintiffs cut off communication with the trafficker. But a video of the compiled images that they sent to the trafficker later appeared on Twitter, which Plaintiffs learned about when the video circulated around their high school. Doe #1 and his mother then began efforts to convince Twitter to remove the video.

Doe #1 filed a complaint through Twitter’s content-reporting interface, and Twitter instructed him to send a copy

of his identification to confirm he was the person in the reported video. He sent his ID and reiterated that the video depicted him and his friend, both minors. Doe #1's mother also reported the posts to Twitter the following day. A few days later, having received nothing but an initial automated response, Doe #1's mother followed up with Twitter to protest its inaction. A couple days after that, Twitter informed Doe #1 that it reviewed the posts, found no policy violations, and would take no further action. In the meantime, the posts received over 150,000 views and 2,000 retweets. Ultimately, nine days after Doe #1's initial report to Twitter, and only at the prompting of the Department of Homeland Security, Twitter removed the posts, suspended the posters' accounts, and reported the content to the National Center for Missing and Exploited Children (NCMEC).

Plaintiffs allege that this episode was not aberrant. Despite developing a variety of tools to monitor and remove content on the platform and having a zero-tolerance policy for child exploitation, Twitter, according to the complaint, is a significant repository for child pornography. Plaintiffs allege Twitter both underutilizes the tools it has developed to curb the spread of this illegal content and has passed on opportunities to develop better tools, despite the inadequacy of its existing infrastructure. Finally, given Twitter's business model, it receives significant advertising revenue from hosting sought-after or popular posts, including those that depict pornographic content featuring minors.

III. Procedural History

Plaintiffs filed a 13-count complaint against Twitter. Twitter moved to dismiss for failure to state a claim, primarily asserting that it is immune from liability under

§ 230. The district court dismissed 12 counts, including, as relevant here, Count 4 (possession and distribution of child pornography in violation of 18 U.S.C. §§ 2252A and 2255), Count 5 (products liability), and Count 8 (negligence per se). The district court concluded that because these counts sought to treat Twitter as a publisher of third-party content—child pornography generally, and the video of Plaintiffs specifically—Twitter was immune under § 230 and Plaintiffs could not state a claim for relief.

At the same time, the district court declined to dismiss Count 2, which alleges that Twitter knowingly benefitted from a sex-trafficking venture in violation of 18 U.S.C. § 1591, giving rise to civil liability under 18 U.S.C. § 1595. Section 230(e)(5)(A), enacted in 2018 as part of the Allow States and Victims to Fight Online Sex Trafficking Act (FOSTA), withdraws § 230(c) immunity for “any claim in a civil action brought under [18 U.S.C. § 1595], if the conduct underlying the claim constitutes a violation of [18 U.S.C. § 1591].” 47 U.S.C. § 230(e)(5)(A). The district court found that Plaintiffs stated such a claim.

The district court certified Twitter’s request for an interlocutory appeal of the decision on Count 2, and we accepted Plaintiffs’ cross-appeal as to Count 4. *See Doe #1 v. Twitter, Inc.*, No. 22-15103, 2023 WL 3220912, at *1 (9th Cir. May 3, 2023). While the first appeal in this case was pending, we addressed for the first time the legal standards governing FOSTA’s § 230(e)(5)(A) immunity exception. *See Does 1–6 v. Reddit, Inc.*, 51 F.4th 1137 (9th Cir. 2022). Following *Reddit*, we reversed as to Count 2 and remanded for the district court to apply our newly articulated standards. *Doe #1*, 2023 WL 3220912, at *2. We also affirmed the dismissal of Count 4. *Id.* Because Plaintiffs’ allegations that Twitter possessed and distributed child pornography “targets

‘activity that can be boiled down to deciding whether to exclude material that third parties seek to post online,’ such activity ‘is perforce immune under section 230.’” *Id.* (quoting *Roommates.com*, 521 F.3d at 1170–71).

On remand, the district court dismissed Count 2 with prejudice because it concluded that Plaintiffs’ claim that Twitter was a beneficiary of sex trafficking did not trigger § 230(e)(5)(A)’s immunity carveout, as interpreted by *Reddit*. Plaintiffs now appeal that decision and the district court’s prior dismissal of their design-defect and negligence per se claims. They also ask us to revisit our prior decision affirming the dismissal of their claim for possession and distribution of child pornography.

DISCUSSION

We review de novo a district court’s dismissal under Federal Rule of Civil Procedure 12(b)(6). *Lathus v. City of Huntington Beach*, 56 F.4th 1238, 1240 (9th Cir. 2023). Twitter bears the burden of establishing that it is immune under § 230. *See Calise*, 103 F.4th at 738 n.1.

As explained above, there are three elements to § 230(c)(1) immunity. *Barnes*, 570 F.3d at 1100–01. The second element, addressing whether the plaintiff seeks to hold the defendant liable “as a publisher or speaker,” is often the most difficult. A defendant is not immune from liability simply because its status as a publisher is a “but-for” cause of a plaintiff’s injuries. *HomeAway.com*, 918 F.3d at 682. Rather, as discussed above, we inspect the alleged legal duty underlying a plaintiff’s claim. First, we examine the “‘right’ from which the duty springs.” *Calise*, 103 F.4th at 742. If it springs “from the defendant’s status as a publisher,” or if the means to avoid liability requires the defendant to act as a

publisher, immunity applies. *Id.* Otherwise, immunity is off the table. *Id.*

I. Beneficiary of Sex Trafficking

As part of the Trafficking Victims Protection Reauthorization Act of 2003 (TVPRA), Congress criminalized sex trafficking, 18 U.S.C. § 1591, and created a civil action for victims to recover damages for criminal violations, *id.* § 1595. Section 1591 punishes anyone who “knowingly” “benefits, financially or by receiving anything of value, from participation in a venture which has engaged in” a sex-trafficking activity. *Id.* § 1591(a). One participates in a sex-trafficking venture by “knowingly assisting, supporting, or facilitating a” prohibited sex-trafficking act, including child sex trafficking. *Id.* § 1591(e)(4). There are two layers of knowledge required: the defendant must “*knowingly* benefit from *knowingly* participating in child sex trafficking.” *Reddit*, 51 F.4th at 1145 (emphasis added).

Plaintiffs seek to hold Twitter liable for knowingly benefitting from a sex-trafficking venture under § 1595 by receiving financial gain or other benefit from maintaining exploitive images of Plaintiffs on its platform. Plaintiffs argue that their TVPRA claim does not treat Twitter as a publisher and, even if it does, FOSTA’s carveout to § 230 immunity applies. These arguments run headlong into our decisions in *Reddit* and *Doe v. Grindr Inc.*, 128 F.4th 1148 (9th Cir. 2025).

A. Does § 230 Immunity Apply?

Plaintiffs contend that Twitter’s duty not to “knowingly benefit” from participation in a sex-trafficking venture does not spring from Twitter’s function as a publisher and that Twitter has multiple means to avoid such benefit that do not

involve monitoring third-party content. *See* 18 U.S.C. §§ 1591, 1595; *Calise*, 103 F.4th at 742.² That is particularly so here, Plaintiffs argue, because Twitter knew that the images at issue were child pornography. Plaintiffs’ thinking is there can be no monitoring duty if the posts are already on Twitter’s radar. But contrary to Plaintiffs’ reasoning, their theory for liability imposes a monitoring obligation. Plaintiffs allege that, given Twitter’s advertising structure and other revenue-generating activities, “[a]s long as content on Twitter’s platform remains live, Twitter monetizes that content.” With that alleged one-to-one relationship between posted content and Twitter monetizing that content, the only way for Twitter to avoid the unlawful benefit from hosting child pornography would be to remove third-party posts—a quintessential publishing activity. *See Roommates.com*, 521 F.3d at 1170–71. As we explained in *Roommates.com*, “any activity that can be boiled down to deciding whether to exclude material that third parties seek to post online is perforce immune under section 230.” *Id.* Accordingly, § 230 immunity applies, absent the FOSTA exception.

² We reject Twitter’s contention that the law of the case and law of the circuit foreclose Plaintiffs’ argument that § 230(c) does not apply. As to law of the case, the applicability of § 230(c) was not one of the limited questions certified by the district court and accepted by this court for interlocutory review, nor was it a necessary predicate to the issue we decided there. *Doe #1*, 2023 WL 3220912, at *1. And regarding law of the circuit, in *Reddit* the parties agreed that § 230(c) applied and so we had no occasion to analyze that question. 51 F.4th at 1141; *see Webster v. Fall*, 266 U.S. 507, 511 (1925) (“Questions which merely lurk in the record, neither brought to the attention of the court nor ruled upon, are not to be considered as having been so decided as to constitute precedents.”).

B. Does FOSTA’s Exception to Immunity Apply?

FOSTA displaces § 230 immunity for civil actions brought under 18 U.S.C. § 1595 “if the conduct underlying the claim” violates 18 U.S.C. § 1591. 47 U.S.C. § 230(e)(5)(A). For that carveout to apply, Plaintiffs “must plausibly allege that [Twitter’s] own conduct violated section 1591.” *Reddit*, 51 F.4th at 1141. It is not enough to allege “[m]ere association with sex traffickers” or “trafficking by the website’s users—without the participation of the website.” *Id.* at 1142, 1145. Here, the district court concluded that Plaintiffs did not sufficiently allege an underlying § 1591 violation because they did not allege that Twitter itself violated § 1591. We agree.

Reddit involved similar allegations. Parents of minors whose pictures were posted on Reddit’s website alleged that Reddit made it dangerously easy for users to share child pornography, highlighted subsections of its platform that contained child pornography to gain advertising revenue, and failed to remove child pornography when it was reported, including by the plaintiffs in that case. *Id.* at 1145. We held that these allegations did not state a § 1591 sex-trafficking violation. *Id.* at 1145–46. Alleging that the defendant “merely turn[ed] a blind eye to the source of their revenue” was insufficient to allege Reddit “actively participated in sex trafficking.” *Id.* at 1145 (citation modified). But that was as far as the plaintiffs’ allegations went.

Plaintiffs argue that Twitter cannot claim that it merely turned a blind eye because Twitter confirmed that it reviewed Plaintiffs’ reports and the flagged posts. While Twitter’s review of the challenged videos that were posted by a trafficker establishes its knowledge that the videos

contained child pornography, Twitter’s failure to respond to demands to remove the videos is not the type of “affirmative conduct” that constitutes “assistance, support, or facilitation” of sex trafficking for which § 1591 attaches criminal (and, correspondingly, civil) liability. *See id.* That is, Twitter did not “actually engage[] in some aspect of the sex trafficking,” as a legal matter, by failing to remove known child pornography from its platform. *Grindr*, 128 F.4th at 1155 (quoting *Reddit*, 51 F.4th at 1145). While we understand the logic of Plaintiffs’ argument that continuing to make available known child pornography is tantamount to facilitating sex trafficking, that reasoning fails under our prior holding that merely turning a blind eye to illegal revenue-generating content does not establish criminal liability under § 1591. *Reddit*, 51 F.4th at 1145.

In any event, Plaintiffs rely on the same causal theory that we held insufficient in *Reddit* and *Grindr*. *See Reddit*, 51 F.4th at 1145–46; *Grindr*, 128 F.4th at 1155. We require “a causal relationship between affirmative conduct furthering the sex-trafficking venture and receipt of a benefit.” *Reddit*, 51 F.4th at 1145 (citation omitted). Generic advertising revenue schemes that apply sitewide do not satisfy that causal demand. *See id.* at 1145–46 (concluding that the allegation “that Reddit makes money from advertising on all popular subreddits” did not establish the requisite causal nexus); *Grindr*, 128 F.4th at 1155–56 (holding that allegations of Grindr “generally receiving advertising revenues” did not “causally connect Grindr’s advertising revenues with any affirmative conduct by Grindr that furthered the sex-trafficking venture”). But that is Plaintiffs’ theory of knowing benefit here: Twitter profits from all the posts on its website, it knew the posts at issue here contained child pornography, and therefore it

knowingly benefited from a child-pornography trafficking venture. Under our precedents, that simply does not suffice to state a claim.³ We therefore affirm the district court's dismissal of Count 2.

II. Design Defect

Plaintiffs allege that Twitter has design defects that violate California's products-liability law. The alleged deficiencies fall into three categories: (1) Twitter lacks a mechanism allowing users to report child pornography easily, (2) Twitter does not block reported child pornography pending review of a complaint, and (3) Twitter's search feature that includes search suggestions and hashtags amplifies the reach of child pornography once this content is posted. We take each alleged defect in turn.

A. Reporting Infrastructure

Plaintiffs allege that Twitter makes it too difficult to report child pornography that is posted on Twitter. For example, someone wishing to report this content cannot use Twitter's "easily-accessible" general report function, but must instead locate a separate form that is unique to reporting child pornography. And that form has

³ Plaintiffs also argue that the district court erred by analyzing only whether Twitter had acted with knowledge that a minor was induced to engage in a commercial sex act when § 1591 also criminalizes acting with "reckless disregard." See 18 U.S.C. § 1591(a)(2). Even if Plaintiffs were correct, this alleged mens rea error would not impact the outcome here because Plaintiffs' pleading deficiencies concerning Twitter's affirmative conduct and causation remain. *Cf. M.H. ex rel. C.H. v. Omegle.com LLC*, 122 F.4th 1266, 1275 (11th Cir. 2024) (characterizing FOSTA's mens rea standard as a "slightly different question" than that addressed in *Reddit*: "to whose conduct does FOSTA refer . . . ?"). Thus, we express no opinion on whether the district court misapplied § 1591's mens rea requirement.

disadvantages: it does not allow a user to report child pornography sent via private messaging, it requires reporters to supply an email address, and it requires a person to have and be logged into a Twitter account. According to Plaintiffs, child-protection watchdogs have faulted Twitter’s process, opining that it lags behind Twitter’s peers.

This aspect of Plaintiffs’ design-defect claim relates solely to product design, and in that way, is analogous to the products-liability claim that we allowed to proceed in *Lemmon v. Snap, Inc.*, 995 F.3d 1085 (9th Cir. 2021). The plaintiffs in *Lemmon* challenged a Snapchat filter that showed the speed a user was traveling, which the plaintiffs alleged encouraged reckless driving. *Id.* at 1088–89. When a young driver used the speed filter shortly before a fatal crash, his parents sued for negligent design. *Id.* at 1088, 1091–92. We allowed the claim to proceed, reasoning that it turned on Snap’s design architecture rather than the publication of any content. *Id.* at 1092–93. We underscored that the plaintiffs’ claim treated Snap as a “product designer” rather than a “publisher or speaker” because “Snap could have satisfied its ‘alleged obligation’ . . . without altering the content that Snapchat’s users generate.” *Id.* at 1092 (citation omitted).

So too here. Twitter could fulfill its purported duty to cure reporting infrastructure deficiencies without monitoring, removing, or in any way engaging with third-party content. *See id.* This claim thus does not seek to hold Twitter responsible as a publisher or speaker. *Id.* Increased removal of third-party content may well be the outcome of a more robust reporting structure. But a claim alleging a duty that does not treat a defendant as a publisher is not barred by § 230, even if that legal duty “might lead a company to respond with monitoring or other publication activities.” *HomeAway.com*, 918 F.3d at 682. “We look instead to what

the duty at issue *actually requires*: specifically, whether the duty would necessarily require an internet company to monitor third-party content.” *Id.* (emphasis added). Here, Twitter’s improvement of its reporting mechanism—for example, by allowing people to report child pornography sent via private messaging—would not necessarily require Twitter to monitor third-party content. Accordingly, Twitter is not immune from liability on this claim.

B. Removal Pending Review

Plaintiffs also complain that Twitter “failed to block reported [child pornography] while [it] was investigated and enabled reported [child pornography] to continue to be massively disseminated.” This product-defect theory is barred by § 230. Plaintiffs are attempting to hold Twitter liable for its failure to remove (even if automatically) harmful third-party content. As we have previously recognized, decisions about what third-party content is disseminated are quintessentially the province of a publisher, triggering § 230 immunity. *See Roommates.com*, 521 F.3d at 1174 (explaining that § 230 was “enacted to protect websites against the evil of liability for failure to remove offensive content.”); *Barnes*, 570 F.3d at 1102 (“[P]ublication involves . . . deciding whether to publish or to withdraw from publication third-party content.”).

C. Amplification

Finally, Plaintiffs allege that Twitter designed an unreasonably dangerous product that allows “search features and hashtags that are specifically for the purpose of searching, finding, and sharing [child pornography].” They allege Twitter owes a duty to design a reasonably safe product that prevents its search feature from responding to hashtags used solely to identify child pornography and from

providing search suggestions that enable users to find such content. As we understand it, Plaintiffs take no issue with the general interaction between the hashtag and search functions when put to lawful ends. For example, to highlight posts that are tagged #ParisOlympics when a user searches “Olympics.” Rather, they allege that sex traffickers create and use certain hashtags to signal child pornography. Plaintiffs allege that Twitter’s search function defectively responds to these nefarious hashtags, aiding consumers of child pornography in finding the illegal content that they are looking for.

We have been skeptical of similar efforts to hold internet content providers liable for developing “content-neutral tools used to facilitate communications.” *Dyroff v. Ultimate Software Grp., Inc.*, 934 F.3d 1093, 1096 (9th Cir. 2019). In *Roommates.com*, we predicted that “ordinary search engines” like Google and Yahoo! would be entitled to § 230 immunity. 521 F.3d at 1167. To permit liability for generic search functions, we explained, “would defeat the purposes of section 230 by swallowing up every bit of the immunity that the section otherwise provides.” *Id.*

Accordingly, in *Dyroff*, we allowed § 230 immunity for a website whose machine-learning algorithms and facially neutral features like interest-driven groups and notifications promoted illegal-drug posts to someone who had previously sought out such content. 934 F.3d at 1095–96, 1098. We concluded that the plaintiff’s claim treated the website as a publisher of third-party content, even if the website’s tools helped disseminate and promote that content. *Id.* at 1098–99; *accord M.P. ex rel. Pinckney v. Meta Platforms Inc.*, 127 F.4th 516, 521, 523–27 (4th Cir. 2025) (barring claims that “attack the manner in which Facebook’s algorithm sorts, arranges, and distributes third-party content”). So, too, when

the victims of terrorist attacks alleged Google’s algorithm promoted terrorist content on YouTube. *See Gonzalez v. Google LLC*, 2 F.4th 871, 881–82, 892–96 (9th Cir. 2021), *vacated*, 598 U.S. 617 (2023) (per curiam) (“[A] website’s use of content-neutral algorithms, without more, does not expose it to liability for content posted by a third-party.”). And in *Grindr*, we held that immunity applied to a claim alleging that a neutral romantic-matching tool was negligently designed because it created opportunities for sexual abuse. 128 F.4th at 1152–53.

Against this backdrop, we hold that Twitter is immune from liability for the alleged third-party abuses of its hashtag and search functions. Distinguishing between innocent #ParisOlympics-type hashtags and the more nefarious ones would require Twitter to act as a publisher. Notwithstanding Plaintiffs’ allegation that “Twitter has the ability to, and in fact does, block certain hashtags,” deciding when to take that step is a publisher decision. *See Roommates.com*, 521 F.3d at 1174 (clarifying that removal decisions are publisher decisions). Otherwise, upon designing neutral hashtag and search tools, Twitter would be required to monitor and act when its users adapt them to illicit ends.

III. Negligence Per Se

Plaintiffs also allege that Twitter was negligent per se because it did not promptly report the exploitive video of them to NCMEC, in violation of 18 U.S.C. § 2258A.⁴ The

⁴ The complaint also alleges negligence per se premised on violations of 18 U.S.C. §§ 1591 and 1595, California Civil Code § 1708.85, and California Penal Code § 311.1. On appeal, Plaintiffs focus exclusively on 18 U.S.C. § 2258A’s reporting requirement.

district court held that § 230 barred this claim, along with Plaintiffs' negligence and gross negligence counts.

At the outset, Twitter argues that Plaintiffs waived their ability to treat negligence per se as “a viable, standalone claim” by telling the district court that it “should treat the negligence per se claim as a support to, and part of” their negligence and gross-negligence counts. To be sure, Plaintiffs could have been clearer in developing their negligence per se claim. But they insisted, in their opposition to Twitter's motion to dismiss, that this claim was “not duplicative” of their negligence and gross-negligence counts. And even if Plaintiffs had waived this claim below, we would exercise our discretion to consider it on appeal. *See Hoffman v. Pulido*, 928 F.3d 1147, 1150 (9th Cir. 2019).

Our waiver rules in this context exist in large part to ensure that the parties, and we, do not sandbag the district court once it has committed to a course of action in response to the issues and arguments the parties presented. *See Whittaker Corp. v. Execuair Corp.*, 953 F.2d 510, 515 (9th Cir. 1992). But here, the district court did not rely on Plaintiffs' supposed concession that Twitter highlights. That statement arose in response to Twitter's argument that negligence per se is not an independent cause of action under California law.⁵ The district court expressly reserved

⁵ Twitter continues to press that argument on appeal, along with its contention that this is not one of the limited circumstances in which California allows recovery for nonfeasance rather than misfeasance. Although “we may affirm based on any ground supported by the record,” *Johnson v. Riverside Healthcare Sys., LP*, 534 F.3d 1116, 1121 (9th Cir. 2008), we typically leave merits arguments for remand when we reverse a district court's conferral of § 230 immunity, e.g., *Lemmon*, 995 F.3d at 1094–95; *Calise*, 103 F.4th at 743 n.5. We follow that usual course here.

judgment on that state-law argument because it concluded that Twitter was immune under § 230.

We reverse that § 230 holding. Under 18 U.S.C. § 2258A(a)(1)(A)(i), electronic-communication services are required to file a report with NCMEC “as soon as reasonably possible after obtaining actual knowledge of” violations of criminal sex-trafficking laws involving child pornography. Twitter argues that its reporting duty arises because its platform allows third parties to upload content. That may well be true. Twitter “is an internet publishing business” and “publishing content is ‘a but-for cause of just about everything’ [Twitter] is involved in.” *Lemmon*, 995 F.3d at 1092–93 (citation omitted). But that is not the test. *Id.* And the facts alleged here, coupled with the statutory “actual knowledge” requirement, separates the duty to report child pornography to NCMEC from Twitter’s role as a publisher.

Plaintiffs do not claim that Twitter must scour its platform for content triggering its NCMEC-reporting duty. They do not even claim that Twitter must review reported child pornography. Rather, they allege that once Twitter has obtained *actual knowledge* of such content, as evidenced by its representation that it had “reviewed the content,” it had a legal duty to promptly report that content to NCMEC. Because that duty neither requires Twitter to monitor content nor take any action associated with publication (e.g., removal) once it learns of the objectionable content, § 230 does not immunize Twitter from Plaintiffs’ negligence per se claim.

IV. Possession and Distribution of Child Pornography

Finally, Plaintiffs appeal the district court’s dismissal of their claim alleging that Twitter possessed and distributed child pornography in violation of 18 U.S.C. § 2255. We have

already affirmed that dismissal. *Doe #1*, 2023 WL 3220912, at *2. That is the end of the matter. *See Gonzalez v. Arizona*, 677 F.3d 383, 389 n.4 (9th Cir. 2012).

CONCLUSION

For the foregoing reasons, we affirm the district court's dismissal of Count 2 (TVPRA), Count 4 (18 U.S.C. § 2255), and Count 5 (California Products Liability) as relates to Plaintiffs' removal and amplification design-defects theories; reverse the district court's dismissal of Count 5 as relates to Plaintiffs' defective-reporting-infrastructure theory and Count 8 (negligence per se); and remand for further proceedings consistent with this opinion.

**AFFIRMED IN PART; REVERSED IN PART;
REMANDED.⁶**

⁶ Each party is to bear its own costs.