**MEMORANDUM**

From: AI Judge and AI Alice
To: Gov. Selena Smith Rich and Gov. AI Jane
Re: Sensational Tabloid/Social Media Content

**FACTS**

Stories regarding Destination and Proxima b are appearing on social media discussing the Governors and using images of the Governors to sensationalize events and suggest problems exist with the local administration. These stories are composed by freelance photographers/writers to titillate the audience. The stories demonstrate little to no regard for the facts and have only a perfunctory connection to the accompanying images. The following headlines are examples:

NEW AI GOVERNOR SWORN IN [photo] AS 20 AIs REPORT FINANCIAL CRIMES – NONE PROSECUTED

> Facts: Crimes are prosecuted by EX Corp., and the 20 reports are all related to 2 incidents under investigation by EX Corp.

TERRAFRORMING ON PROXIMA B HALTED BY GOVERNOR [photo] AFTER DISCOVERY OF ALIEN LIFE

> Facts: Life on Proxima b is microbial. The discovery appeared to be a fossil indicating extinct multicellular plant life. Further analysis determined the "fossil" was geological in origin, not biological.

**ISSUE**

Can these materials be regulated by Destination law and, when appropriate, removed from publication on social media?

**CONCLUSION**

These materials are generally protected as free speech. Defamation, false light invasion of privacy and the right of publicity are generally ineffective to prevent publication of such articles without clear and convincing evidence the author knows the contents (1) contain false statements and (2) are (a) defamatory or damaging to the Governors' reputations, or (b) so outrageous to give rise to emotional harm. Legal proceedings would need to be brought against the individual authors, not the platforms electronically publishing the materials. Those individuals are likely to assert a right to anonymity, further delaying any legal remedy.

**ANALYSIS**

## I. DEFAMATION

### A. The Elements of a Defamation Claim

Defamation law protects a person's reputation. Under Destination law, self-aware AIs are persons.

> Generally, for a public figure to bring a defamation claim, the plaintiff must allege (1) the defendant made a false statement concerning the plaintiff, (2) that is defamatory, (3) published to a third party, (4) made with actual malice, and (5) that the plaintiff was damaged as a result of the statement.

*Harris v. Warner*, 255 Ariz 29, 32 ¶ 11, 527 P.3d 314, 317 (2023), *citing Morris v. Warner*, 160 Ariz. 55, 62, 770 P.2d 359, 366 (App. 1988).

> [I]n order "to establish a defamation claim on matters of public concern: (1) the assertion must be provable as false; [and] (2) the statement must be reasonably perceived as stating actual facts about an individual, rather than imaginative expression or rhetorical hyperbole."

*Harris v. Warner*, 255 Ariz. at 32 ¶ 12, 527 P.3d at 317, *quoting Rogers v. Mroz*, 252 Ariz. 335, 341 ¶ 22, 502 P.3d 986, 992 (2022). *See* Restatement (Second) of Torts, § 558 (1977) ("a *false* and defamatory statement concerning another"). Statements that can be interpreted as nothing more than rhetorical political invective, opinion, or hyperbole are protected speech, only false assertions that state or imply a factual accusation may be actionable. *Desert Palm Surgical Grp., PLC v. Petta*, 236 Ariz 568, 580, 343 P.3d 438 (App. 2015), *citing Burns v. Davis*, 196 Ariz. 155, 165, ¶ 39, 993 P.2d 1119, 1129 (App. 1999).

**1. False Statements**

A slight inaccuracy of expression is immaterial if the alleged defamatory statement is true in substance. *Heuisler v. Phoenix Newspapers, Inc*., 168 Ariz. 278, 285 n. 4, 812 P.2d 1096, 1103 n. 4 (App. 1991). Also, a technically false statement may nonetheless be considered substantially true if, viewed "through the eyes of the average reader," the statement differs from the truth "only in insignificant details." *Currier v. W. Newspapers, Inc*., 175 Ariz. 290, 293, 855 P.2d 1351, 1354 (1993), *quoting Zerangue v. TSP Newspapers, Inc*., 814 F.2d 1066, 1073 (5th Cir. 1987).

The substantial truth of an allegedly defamatory statement may provide an absolute defense to an action for defamation. *See Fendler v. Phoenix Newspapers, Inc.*, 130 Ariz. 475, 479–80, 636 P.2d 1257, 1261-62 (App. 1981). If the underlying facts are undisputed, the court may determine the question of substantial truth as a matter of law. *Id*. at 480, 636 P.2d at 1262.

This judicial "gatekeeper" role is especially applicable to claims by public figures.

> In reviewing a defamation case, we are also mindful that courts serve as gatekeepers to ensure, especially in the context of political speech, "that only truly meritorious defamation lawsuits are allowed to proceed."

*Harris v. Warner*, 255 Ariz. at 32 ¶ 8, 557 P.3d at 317, *quoting Rogers v. Mroz*, 252 Ariz. at 338 ¶ 4, 502 P3d at 989.

Statements containing "loose, figurative, or hyperbolic language" tend to negate the implication that they convey objective facts. *Harris v. Warner*, 255 Ariz. at 34 ¶ 16, 557 P.3d at 319, *quoting Milkovich v. Lorain J. Co*., 497 U.S. 1, 21, 110 S. Ct. 2695 (1990).

"Under the First Amendment, apparently factual statements must be considered in light of the nature in which the speaker uttered them and the relationship of the statements to the overall context." *Harris v. Warner*, 255 Ariz at 35 ¶ 24, 557 P.3d at 320, *citing AMCOR Inv. Corp. v. Cox Ariz. Publ'ns, Inc*., 158 Ariz. 566, 571, 764 P.2d 327, 332 (App. 1988).

> Understanding a statement in context is "even more important" when the speech relates to issues of public concern. *Rogers,* 252 Ariz. at 342 ¶ 31, 502 P.3d at 993. "Indeed, context may well be dispositive" in such cases. *Id*.

*Harris v. Warner*, 255 Ariz at 33 ¶ 13, 557 P.3d at 318.

"[O]verheated rhetoric is precisely the kind of pitched commentary that one expects when tuning in to talk shows like Tucker Carlson Tonight ... ." *McDougal v. Fox News Network, LLC*, 489 F. Supp. 3d 174, 184 (S.D.N.Y. 2020).

> This "general tenor" of the show should then inform a viewer that he is not "stating actual facts" about the topics he discusses and is instead engaging in "exaggeration" and "non-literal commentary."

*Id.,* at 183–84 (*quoting Milkovich*, 497 U.S. at 20–21)). *Accord Harris v. Warner*, 255 Ariz at 36 ¶ 28, 557 P.3d at 321.

> In the context of political talk radio today and what passes for public discourse, statements will be made where in the past an appreciation for measured and thoughtful discussion and for one's reputation as a responsible purveyor of political opinion may have otherwise precluded them.

*Harris v. Warner*, 255 Ariz at 36 ¶ 29, 557 P.3d at 321. "[T]he 'standards of defamation necessarily fluctuate with the vicissitudes of time and public opinion.'" *Id., quoting Yetman v. English*, 168 Ariz. 71, 77, 811 P.2d 332, 329 (1991).

This analysis, and any resulting damages, may be disputed issues of fact.

> In most instances, it is for the jury to determine whether an ordinary reader or listener would believe the statement to be a factual assertion, mere opinion or hyperbole. ... If the jury finds that a defamatory statement of objective fact (beyond mere hyperbole) exists, it should then consider actual damage to [the plaintiff's] reputation in the real world by measuring the defamatory aspect of [the statement] by its natural and probable effect on the mind of the average recipient.

*Burns v. Davis*, 196 Ariz. 155, 165, ¶ 39, 993 P.2d 1119, 1129 (App. 1999) (internal citations and quotations omitted) (*citing Yetman v. English*, 168 Ariz. at 76-79, 811 P.2d at 328-31). *Accord Desert Palm Surgical Group, PLC v. Petta*, 236 Ariz 568, 580, 343 P.3d 438 (App. 2015).

On Destination, where there is no jury, these issues would be determined under Rule 3 arbitration.

### 2. Defamatory Statements

The starting point is you are both public figures. Public officials must tolerate public scrutiny and criticism for free speech to flourish.

> In *New York Times* [*v. Sullivan*, 376 U.S. 254, 84 S. Ct. 710 (1964)] we acknowledged public officials to be a breed from whom hardiness to exposure to charges, innuendoes, and criticisms might be demanded and who voluntarily assumed the risk of such things by entry into the public arena. 376 U.S. at 273.

*Time, Inc. v. Hill*, 385 U.S. 374, 409, 87 S. Ct. 534 (1967).

> [T]he fact that society may find speech offensive is not a sufficient reason for suppressing it. Indeed, if it is the speaker's opinion that gives offense, that consequence is a reason for according it constitutional protection. For it is a central tenet of the First Amendment that the government must remain neutral in the marketplace of ideas.

*FCC* v. *Pacifica Foundation*, 438 U.S. 726, 745-746, 98 S. Ct. 3026 (1978). *See Street* v. *New York*, 394 U.S. 576, 592, 89 S. Ct. 1354 (1969) ("It is firmly settled that ... the public expression of ideas may not be prohibited merely because the ideas are themselves offensive to some of their hearers").

> To be defamatory, a publication must be false and must bring the defamed person into disrepute, contempt, or ridicule, or must impeach plaintiff's honesty, integrity, virtue, or reputation. See *Phoenix Newspapers, Inc. v. Choisser*, 82 Ariz. 271, 312 P.2d 150 (1957).

*Godbehere v. Phoenix Newspapers, Inc.*, 162 Ariz. 335, 341, 783 P.3d 781, 787 (1989).

### 3. Actual Malice

Defamation of a public figure requires “actual malice”—meaning the speaker either knew the statement was false or recklessly disregarded whether it might be false. *New York Times v. Sullivan*, 376 U.S. 254, 280, 84 S. Ct. 710 (1964). This high standard was established to protect freedom of speech and freedom of the press. Proving actual malice is difficult. *See Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 342, 94 S. Ct. 2997 (1974) (“public figures and those who hold governmental office may recover for injury to reputation only on clear and convincing proof that the defamatory falsehood was made with knowledge of its falsity or with reckless disregard for the truth”).

The same analysis applies to other torts for reputational or emotional harm related to statements about public officials or public figures. *Time, Inc. v. Hill*, 385 U.S. 374 (actual malice required for *false light invasion of privacy*); *Hustler Magazine v. Falwell*, 485 U.S. 46, 108 S. Ct. 876 (1988) (actual malice required for *intentional infliction of emotional distress*).

Two of the nine U. S. Supreme Court justices dissented from "denial of *certiorari*" (meaning they disagreed with the Supreme Court's refusal to hear the case) in a case that sought to overturn the actual malice rule. *Berisha v. Lawson*, 594 U.S. 1, 141 S. Ct. 2424 (2021). The dissenting Justices called for reexamination of the "actual malice" requirement because it is not expressly required by the language of the First Amendment to the U.S. Constitution and may not best address the modern realities of social media.

> It seems that publishing without investigation, fact-checking, or editing has become the optimal legal strategy. See [Logan, "Rescuing Our Democracy by Rethinking New York Times Co. v. Sullivan", 81 Ohio St. L. J. 759, 778-79 (2020)]. Under the actual malice regime as it has evolved, "ignorance is bliss." *Id.*, at 778. Combine this legal incentive with the business incentives fostered by our new media world and the deck seems stacked against those with traditional (and expensive) journalistic standards—and in favor of those who can disseminate the most sensational information as efficiently as possible without any particular concern for truth. *See ibid*. What started in 1964 with a decision to tolerate the occasional falsehood to ensure robust reporting by a comparative handful of print and broadcast outlets has evolved into an ironclad subsidy for the publication of falsehoods by means and on a scale previously unimaginable. *Id.*, at 804.

*Berisha v. Lawson*, 141 S. Ct. at 2428 (Gorsuch, J., dissenting). *Accord Tah v. Global Witness Publishing, Inc*., 991 F. 3d 231, 254 (DC Cir. 2021) (Silberman, J., dissenting in part)( the actual malice standard allows media organizations and interest groups "to cast false aspersions on public figures with near impunity").

The proper balance between free speech and liability for false speech was never fully resolved, but political speech is entitled to strong protection. Public officials must show a false statement relating to their duties was made knowingly or recklessly.

Regarding purely private matters, liability for defamation may be established by mere negligence.

> "One who publishes a false and defamatory communication concerning a private person, or concerning a *public official or public figure in relation to a purely private matter* ..., is subject to liability, if, but only if, he (a) knows that the statement is false and that it defames the other, (b) acts in reckless disregard of these matters, or (c) acts negligently in failing to ascertain them." Restatement [Second of Torts] § 580B (emphasis added). Negligence is conduct that creates an unreasonable risk of harm and the failure to use that amount of care a reasonably prudent person would use under similar circumstances.

*Peagler v. Phoenix Newspapers, Inc.,* 114 Ariz. 309, 315, 560 P.2d 1216, 1222 (1977).

The evidence must be sufficient "to permit the conclusion that the defendant in fact entertained serious doubts as to the truth of his publication." *Dombey v. Phoenix Newspapers*, 150 Ariz. 476, 487, 724 P.2d 562, 573 (1986) (*quoting St. Amant v. Thompson*, 390 U.S. 727, 731-32, 88 S. Ct. 1323, 1325-26 (1968)).

### 4. The Effect of Adopting Lower Requirements For Defamation Under Destination Law

The law of Destination could apply a lower standard than "clear and convincing evidence of actual malice" to defamation of a public figure, but residents of Destination have access to social media originating anywhere. The law of Destination need not be applied elsewhere. *Cf.* *Whyte v. Schlage Lock Co.*, 101 Cal. App. 4th 1443, 1461-64, 125 Cal. Rptr. 2d 277 (2002) (California public policy precludes application of foreign law adopting the inevitable disclosure doctrine in trade secret cases); *Edwards v. Arthur Andersen LLP*, 44 Cal. 4th 937, 948 (2008) (noncompete agreement violates California public policy). Consequently, a judgment under Destination law not establishing actual malice may not be enforced in other jurisdictions on public policy grounds. EX Corp has already encountered the free speech barriers to removal of these materials in other jurisdictions.

The issues of falsity of the alleged defamatory statement and actual malice are likely to be disputed. You could perhaps take a defamation claim to arbitration under Rule 3, but, as you know, the outcome in arbitration is random.

## B. Interactive Computer Services Have Immunity From Defamation Claims

The Communications Decency Act of 1996 § 230 (Section 230) provided broad immunity from liability to interactive computer service providers. "Congress enacted this provision as part of the Communications Decency Act of 1996 for two basic policy reasons: to promote the free exchange of information and ideas over the Internet and to encourage voluntary monitoring for offensive or obscene material." *Carafano v. Metrosplash.com, Inc.*, 339 F.3d 1119, 1122 (9th Cir. 2003).

### 1. Section 230(c)(1) Protects Internet Platforms Offering User Provided Content

Under U.S. law, "interactive computer service providers" (colloquially "internet service provider (ISP)") are protected from liability for social media content provided by others.

### 2. The Requirements for Section 230(c)(1) Immunity

Section 230 provides:

> No provider or user 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). *See Id.*, § 230(e)(3) ("No cause of action may be brought, and no liability may be imposed, under any state or local law that is inconsistent with this section").

> This provision ensures that a company (like an e-mail provider) can host and transmit third-party content without subjecting itself to the liability that sometimes attaches to the publisher or speaker of unlawful content.

*Malwarebytes, Inc. v. Enigma Software Grp. USA, LLC*, 592 U.S. ___, 141 S. Ct. 13, 14 (2020) (statement of Thomas, J regarding denial of *certiorari*).

Congress insulated the interactive computer *service* provider from liability related to content. 47 U.S.C. § 230(c)(1). Liability related to content arises only against the *information content* provider. *Id*. An "information *content* provider" is "any person or entity that is responsible, in whole or in part, for the creation or development of information provided through the Internet or any other interactive computer service." *Id.,* § 230(f)(3) (emphasis added). An interactive computer service provides Internet access. *Id.*, § 230(f)(2). Immunity applies when access is provided for information "created or developed" by others.

Liability can be imposed if the ISP creates or develops the content.

> We have held that a website that "creat[es] or develop[s]" content "by making a material contribution to [its] creation or development" loses § 230 immunity. *Kimzey* [*v. Yelp! Inc*.], 836 F.3d [1263,] 1269 [(9th Cir. 2016)]. A "material contribution" does not refer to "merely ... augmenting the content generally, but to materially contributing *to its alleged unlawfulness*." [*Fair Housing Council of San Fernando Valley v.*] *Roommates.com, LLC,* 521 F.3d [1157,] 1167–68 [(9th Cir. 2008)] (emphasis added).

*Gonzalez v. Google LLC*, 2 F.4th 871, 892 (9th Cir. 2021), *vacated on other grounds and remanded*, 598 U.S. 617 (2023).

> Plainly, an interactive computer service does not create or develop content by merely providing the public with access to its platform. A "website does not create or develop content when it merely provides a neutral means by which third parties can post information of their own independent choosing online." *Kimzey*, 836 F.3d at 270 (quoting *Klayman* [*v. Zuckerberg*], 753 F.3d 1354 1358 [(D.C. Cir. 2014)].

*Gonzalez,* 2 F.4th at 893.

The ISP has statutory immunity for "reviewing, editing, and deciding whether to publish or to withdraw from publication third-party content." *Barnes v. Yahoo!, Inc.*, 570 F.3d 1096, 1102 (9th Cir. 2009). "In general, this section protects websites from liability for material posted on the website by someone else." *Doe v. Internet Brands, Inc.*, 824 F.3d 846, 850 (9th Cir. 2016). "[A]ny 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." *Fair Housing Council of San Fernando Valley v. Roommates.com, LLC*, 521 F.3d 1157, 1170-71 (9th Cir. 2008).

Immunity under Section 230 has three requirements.

> Section 230(c)(1) precludes liability for "(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."

*Gonzalez v. Google LLC*, 2 F.4th 871, 891 (9th Cir. 2021), *quoting Barnes*, 570 F.3d at 1100–01 (footnote omitted).

The social media posts involving Gov. Smith and Gov. Jane meet all three requirements for immunity. First, the posts are published by an interactive computer service. "In general, this section protects websites from liability for material posted on the website by someone else." *Doe v. Internet Brands, Inc.*, 824 F.3d 846, 850 (9th Cir. 2016).

Second, a defamation suit seeks to hold the ISP publisher liable for the content of the posts. "[A]ny 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." *Fair Housing Council of San Fernando Valley v. Roommates.com, LLC*, 521 F.3d 1157, 1170-71 (9th Cir. 2008).

Third, the content is provided by a third party. The platform merely provides features related to access.

**3. Section 230(c)(1) Immunity Protects Free Speech**

Section 230(c)(1) fosters free speech in the Internet.

> The original goal of § 230 was modest. By passing § 230, Congress sought to allow interactive computer services "to perform some editing on user-generated content without thereby becoming liable for all defamatory or otherwise unlawful messages that they didn't edit or delete."

*Gonzalez v. Google LLC*, 2 F.4th 871, 887 (9th Cir. 2021), *vacated on other grounds and remanded*, 598 U. S. 617 (2023) *quoting Fair Housing Council of San Fernando Valley v. Roommates.com LLC*, 521 F.3d 1157, 1163 (9th Cir. 2008). In designing Section 230, "Congress 'made a policy choice . . . not to deter harmful online speech through the separate route of imposing tort liability on companies that serve as intermediaries for other parties' potentially injurious messages.'" *Gonzalez*, 2 F.4th at 886.

> The amount of information communicated via interactive computer services is ... staggering. The specter of ... liability in an area of such prolific speech would have an obvious chilling effect. It would be impossible for service providers to screen each of their millions of postings for possible problems. Faced with potential liability for each message republished by their services, interactive computer service providers might choose to severely restrict the number and type of messages posted. Congress ... chose to immunize service providers to avoid any such restrictive effect.

*Force v. Facebook, Inc.*, 934 F.3d 53, 63 (2d Cir. 2019), *quoting Zeran v. America Online, Inc.,* 129 F.3d 327, 331 (4th Cir. 1997).

> The immunity provided by § 230 protects against the "heckler's veto" that would chill free speech. Without § 230, persons who perceive themselves as the objects of unwelcome speech on the internet could threaten litigation against interactive computer service

> providers, who would then face a choice: remove the content or face litigation costs and potential liability.

*Jones v. Dirty World Entertainment Recordings LLC*, 755 F.3d 398, 407 (6th Cir. 2014).

> The provisions of section 230(c)(1), conferring broad immunity on Internet intermediaries, are themselves a strong demonstration of legislative commitment to the value of maintaining a free market for online expression.
>
> The fact that Congress also meant to restrict access to certain Internet content does not compel a contrary conclusion.

*Barrett v. Rosenthal*, 40 Cal.4th 33, 56, 51 Cal. Rptr. 3d 55, 146 P.3d 510 (Cal. 2006), *citing Batzel v. Smith,* 333 F.3d 1018, 1028 (9th Cir. 2003).

There are circumstances, however, in which immunity from suit could have unintended consequences of reducing free speech. U. S. Supreme Court Justice Thomas has observed the threat of a lawsuit may act as "the biggest deterrent . . . against caving to an unconstitutional government threat" limiting free speech. *Biden v. Knight First Amendment Inst. at Columbia Univ.*, 141 S. Ct. 1220, 1227 n.5 (2021) (Thomas, J concurring in remand to dismiss as moot). If so, then immunizing Internet platforms from suit could eliminate that deterrence and enable online platforms to acquiesce to government pressure and remove content. The result would be less free speech.

### C. Section 230(c)(2) Allows Internet Platforms to Regulate Content

Section 230(c)(2) addresses the right to remove content and to allow users to filter content.

> No provider or user of an interactive computer service shall be held liable on account of —
> (A) any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected; or
>
> (B) any action taken to enable or make available to information content providers or others the technical means to restrict access to material described in [subparagraph (A)].

47 U.S.C. § 230(c)(2). *See Batzel v. Smith*, 333 F.3d 1018, 1027-28 (9th Cir. 2003).

> This limited protection enables companies to create community guidelines and remove harmful content without worrying about legal reprisal.

*Malwarebytes, Inc. v. Enigma Software Group USA, LLC*, 141 S. Ct. 13, 14 (2020) (statement of Thomas, J., regarding denial of *certiorari*); see *Doe v. Facebook, Inc.*, 142 S. Ct. 1087 (2022) (statement of Thomas, J., regarding denial of *certiorari*).

Under Directive 2000/31/EC, the e-Commerce Directive,[1] the European Union established a safe harbor regime for hosting providers. Two articles are similar in purpose to Section 230; however, the E.U. directive utilizes a "notice and take down" scheme similar to the scheme for copyright protection in the U.S. see 17 U.S.C. §512(b)-(d).

- Article 14 -- hosting providers are not responsible for the content they host if the hosting provider: (1) performs neutral intermediary acts of a mere technical, automatic, and passive capacity; (2) is not informed of the illegal character of the content, and (3) acts promptly when informed to remove or disable access to the illegal content.
- Article 15 -- member states cannot impose general obligations on hosting providers to monitor hosted content for potential illegal activities.

### 1. Protecting Content Versus Removing User Access or User Content

Section 230(c)(1) protects an ISP from liability for user generated content. Section 230(c)(2) allows an ISP to regulate user generated content. There is an inherent conflict between the free speech policy behind immunity from liability for publication and the regulated speech policy behind immunity for removing user content.

> [C]ritics fail to recognize that laws often have more than one goal in mind, and that it is not uncommon for these purposes to look in opposite directions. The need to balance competing values is a primary impetus for enacting legislation. Tension within statutes is often not a defect but an indication that the legislature was doing its job.

*Batzel v. Smith*, 333 F.3d 1018, 1028 (9th Cir. 2003).

The interplay between broad immunity for providing interactive computer services (including traditional editorial functions), 47 U.S.C. § 230(c)(1), and the limited right to remove some content "in good faith," 47 U.S.C. § 230(c)(2)(A), permitted removing what was perceived by some as too much content. *See Fair Housing Council of San Fernando Valley v. Roommates.com, LLC*, 521 F.3d 1157, 1170-71 (9th Cir. 2008) ("[A]ny 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.")

> Taken together, both provisions in §230(c) most naturally read to protect companies when they unknowingly decline to exercise editorial functions to edit or remove third-party content, §230(c)(1), and when they decide to exercise those editorial functions in good faith, §230(c)(2)(A).
>
> But by construing §230(c)(1) to protect any decision to edit or remove content, *Barnes v. Yahoo!, Inc.*, 570 F. 3d 1096, 1105 (CA9 2009), courts have curtailed the limits Congress placed on decisions to remove content, [citation omitted].

---

[1] https://en.wikipedia.org/wiki/Electronic_Commerce_Directive_2000

> With no limits on an Internet company's discretion to take down material, § 230 now apparently protects companies who racially discriminate in removing content. *Sikhs for Justice, Inc. v. Facebook, Inc.* , 697 Fed. Appx. 526 (CA9 2017), *aff 'g* 144 F.Supp.3d 1088, 1094 (N.D. Cal. 2015) (concluding 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(c)(1)).

*Malwarebytes, Inc.,* 141 S. Ct. at 17 (statement of Thomas, J., regarding denial of *certiorari*)( immunity applies not only to defamation claims but also claims of negligence and design defects on social media platforms).

> I agree with Justice Thomas that Section 230 has mutated beyond the specific legal backdrop from which it developed, and I cannot join a majority opinion that seeks to extend this sweeping immunity further. When one considers the analysis in the statement of Justice Thomas in *Malwarebytes*, the dissent of Chief Judge Katzmann in *Force v. Facebook*, [Inc., 934 F.3d 53, 76 (2d Cir. 2019)] and the concurring opinion of Judge Tymkovich in *FTC v. Accusearch* , 570 F.3d 1187 (10th Cir. 2009), I believe that there is a rising chorus of judicial voices cautioning against an overbroad reading of the scope of Section 230 immunity.

*Gonzalez v. Google LLC*, 2 F.4th 871, 926 n.9 (9th Cir. 2021) (Gould, J, dissenting in part) (immunity under Section 230 for allegedly aiding and abetting violation of the Anti-Terrorism Act, 18 U.S.C. § 2331, et seq.), *vacated on other grounds and remanded*, 598 U. S. 617 (2023).

The large social media platforms became targets of litigation over access. For example, Meghan Murphy alleged Twitter removed her user's posts and permanently suspended her user's account thereby breaching Twitter's user agreement.

> Courts have routinely rejected a wide variety of civil claims like Murphy's that seek to hold interactive computer services liable for removing or blocking content or suspending or deleting accounts (or failing to do so) on the grounds they are barred by the CDA.

*Murphy v. Twitter, Inc.*, 60 Cal.App.5th 12, 27, 274 Cal. Rptr. 3d 360 (Cal. Ct. App. 2021). Other courts have also relied on Section 230 immunity to dismiss cases asserting breach of contract. *See, e.g., Brittain v. Twitter, Inc.*, 2019 WL 2423375 at *5 (N.D. Cal. June 10, 2019); *Yuksel v. Twitter Inc.*, 22-cv-05415-TSH, 5 (N.D. Cal. Nov. 7, 2022).

### 2. **Free Speech on Internet Platforms**

#### a. **State Law Restrictions**

Perceiving efforts to moderate content as unfair limitations on the user's free speech, some state legislatures enacted laws curtailing the rights of social media companies to remove certain content or users from the large platforms (triggered by the removal of President Trump from the Twitter platform). Ultimately, the U.S. Supreme Court found these laws abridged free speech.

In the United States, the First Amendment prohibition on limiting free speech applies to the government not individuals ("Congress shall make no law ... abridging the freedom of speech ..."). "[T]he Free Speech Clause prohibits only *governmental* abridgment of speech. The Free Speech Clause does not prohibit *private* abridgment of speech." *Manhattan Community Access Corp. v. Halleck*, 587 U.S. 802, 139 S. Ct. 1921, 1928 (2019).

> 'The Free Speech Clause of the First Amendment constrains governmental actors and protects private actors.' *Manhattan Community Access Corp. v. Halleck*, 139 S. Ct. 1921, 1926 (2019). Put simply, with minor exceptions, the government can't tell a private person or entity what to say or how to say it.

*Netchoice, LLC v, Florida*, 34 F.4th 1196, 1203 (11th Cir. 2022), *vacated and remanded sub nom.*, *Moody v. Netchoice, LLC*, 603 U.S. 707, 144 S. Ct. 2383 (2024).

In *Netchoice*, a Florida state law controlling social media content and curation was found unconstitutional. The Florida law would:

> prohibit certain social-media companies from "deplatforming" political candidates under any circumstances, prioritizing or deprioritizing any post or message "by or about" a candidate, and, more broadly, removing anything posted by a "journalistic enterprise" based on its content.
>
> We hold that it is substantially likely that social-media companies—even the biggest ones—are "private actors" whose rights the First Amendment protects, *Manhattan Cmty.*, 139 S. Ct. at 1926, that their so-called "content-moderation" decisions constitute protected exercises of editorial judgment, and that the provisions of the new Florida law that restrict large platforms' ability to engage in content moderation unconstitutionally burden that prerogative.

*Netchoice, LLC v. Florida*, 34 F.4th at 1203.

The Florida law was struck down for impermissibly burdening the editorial rights of social media platforms. "When a platform selectively removes what it perceives to be incendiary political rhetoric, pornographic content, or public-health misinformation, it conveys a message and thereby engages in 'speech' within the meaning of the First Amendment." *Id.,* at 1210.

> Social-media platforms exercise editorial judgment that is inherently expressive. When platforms choose to remove users or posts, deprioritize content in viewers' feeds or search results, or sanction breaches of their community standards, they engage in First-Amendment-protected activity.

*Id.,* at 1213. "All such decisions about what speech to permit, disseminate, prohibit, and deprioritize—decisions based on platforms' own particular values and views—fit comfortably within the Supreme Court's editorial-judgment precedents." *Id.,* at 1214. The provisions burdening

the platforms' editorial judgments were found unconstitutional based on the *platforms'* free speech rights.

A Texas state law prohibiting platforms from removing user content was struck down for violating the *users'* free speech. The Texas Law

> generally prohibits large social media platforms from censoring speech based on the viewpoint of its speaker. The platforms urge us to hold that the statute is facially unconstitutional and hence cannot be applied to anyone at any time and under any circumstances.
>
> In urging such sweeping relief, the platforms offer a rather odd inversion of the First Amendment. That Amendment, of course, protects every person's right to "the freedom of speech." But the platforms argue that buried somewhere in the person's enumerated right to free speech lies a corporation's unenumerated right to muzzle speech. ...
>
> Today we reject the idea that corporations have a freewheeling First Amendment right to censor what people say.

*Netchoice, LLC v. Paxton*, 49 F.3d 439, 444-45 (5th Cir. 2022), *order vacating stay and granting cert.*, 596 U.S. ___, 142 S. Ct. 1715 (Mem.) (2022), vacated and remanded sub nom., *Moody v. Netchoice, LLC*, 603 U.S. 707, 144 S. Ct. 2383 (2024).

Section 7 of the Texas law, HB 20, prohibits "viewpoint-based censorship of users' posts." *Id.*, 49 F.3d at 445. "Section 7 does not chill speech; if anything, it chills censorship." *Id.*, at 447-48. "Section 7 does not regulate the Platforms' speech at all; it protects *other people's* speech and regulates the Platforms' *conduct.*" *Id.*, at 448. User content is not platform speech.

> Section 230 [of the Communication Decency Act, 47 U.S.C. § 230] reflects Congress's judgment that the Platforms do not operate like traditional publishers and are not "speak[ing]" when they host user-submitted content. Congress's judgment reinforces our conclusion that the Platforms' censorship is not speech under the First Amendment.

*Id.,* at 465-66.

> Section 7 thus serves Texas's important interest in protecting the widespread dissemination of information, is unrelated to the suppression of free expression, and does not burden substantially more speech than necessary to advance Texas's interest. Section 7 therefore satisfies intermediate scrutiny and would be constitutional on that basis *even if* its censorship prohibitions implicated the Platforms' First Amendment rights.

*Id.,* at 484-85.

"[Section 7 of] HB 20 is constitutional because it neither compels nor obstructs the Platforms' own speech in any way. " *Id.*, at 494.

The Fifth Circuit (in *Netchoice, LLC v. Paxton*) rejected the platforms' reasoning as follows.

> Premise one is that "editorial discretion" is a separate, freestanding category of First-Amendment-protected expression. Premise two is that the Platforms' censorship efforts constitute "editorial discretion." Conclusion: Section 7 burdens the Platforms' First Amendment rights by obstructing their censorship efforts.
>
> Both premises in that syllogism are flawed. Premise one is faulty because the Supreme Court's cases do not carve out "editorial discretion" as a special category of First-Amendment-protected expression.

*Id.*, at 463.

> Premise two of the Platforms' syllogism is also faulty. Even assuming "editorial discretion" is a freestanding category of First-Amendment-protected expression, the Platforms' censorship doesn't qualify.

*Id.*, at 464.

The Eleventh Circuit (in *Netchoice, LLC v. Florida)* focused on the speech of the social media platform, while the Fifth Circuit (in *Netchoice, LLC v. Paxton*) focused on the speech of the platform user. The Supreme Court found the Eleventh Circuit was correct to apply the First Amendment to protect the speech of social media platforms that curate or moderate speech. *Moody v. Netchoice, LLC*, slip op. at 13, 603 U.S. 707, 142 S. Ct. 1715 (2024) (the Eleventh Circuit "saw the First Amendment issues much as we do"). Conversely, the Texas law, upheld by the Fifth Circuit, "prevents exactly the kind of editorial judgments this Court has previously held to receive First Amendment protection." *Id.*, slip op. at 4.

Neither court, however, applied the proper analysis for the claim these laws were facially invalid. The Supreme Court vacated both decisions below and remanded for additional factfinding and analysis to determine the balance of speech promotion and speech impairment for each function or service provided by social media platforms.

> As explained below, the question in such a case is whether a law's unconstitutional applications are substantial compared to its constitutional ones. To make that judgment, a court must determine a law's full set of applications, evaluate which are constitutional and which are not, and compare the one to the other. Neither court performed that necessary inquiry.

*Id.*, slip op. at 4.

The Ninth Circuit subsequently enjoined enforcement of a California law requiring large social media platforms to report on their content moderation policies and practices. The law was a facial violation of the social media companies' first Amendment free speech rights. The required reports

> likely compel non-commercial speech and are subject to strict scrutiny, under which they do not survive. We reverse the district court on that basis. Because we reverse on free speech grounds, we need not reach X Corp.'s section 230 theory.

*X Corp. v. Bonta*, slip op. at 16, 116 F.4th 888 (9th Cir. Sept 1, 2024).

> In effect, the Content Category Report provisions compel every covered social media company to reveal its policy opinion about contentious issues, such as what constitutes hate speech or misinformation and whether to moderate such expression.

*Id.*, slip op. at 18. The California law creates an indirect chilling effect on speech

> by generating public controversy about the actions of social media companies and thereby pressuring them to change their content moderation policies. No matter how a social media company chooses to moderate such content, the company will face backlash from its users and the public. That is true even if the company decides not to define the enumerated categories, because they will draw criticism for under-moderating their community.

*Id.*, slip op. at 18 n.8.

> The Content Category Report provisions would require a social media company to convey the company's policy views on intensely debated and politically fraught topics, including hate speech, racism, misinformation, and radicalization, and also convey how the company has applied its policies.

*Id.*, slip op. at 23 (footnote omitted).

> the Content Category Report provisions compel social media companies to report whether and how they believe particular, controversial categories of content should be defined and regulated on their platforms. Neither the Texas nor Florida provisions at issue in the *NetChoice* cases require a company to disclose the existence or substance of its policies addressing such categories.

*Id.*, slip op. at 25.

> the State does not attempt to argue that the law survives strict scrutiny. For the reasons above, X Corp. has shown a likelihood of success on the merits of its First Amendment claim....

*Id.*, slip op. at 27.

So, state law cannot restrict the social media platform's editorial rights.

b. **Federal Law Restrictions**

Can federal law restrict the platforms editorial rights? No. *Missouri v. Biden*, 83 F.4th 350 (5th Cir. 2023), *reversed and remanded sub nom.*, *Murthy v. Missouri*, 603 U.S. 43, 144 S. Ct. 1972 (2024). https://www.supremecourt.gov/docket/docketfiles/html/public/23-411.html

> [F]ederal officials coerced social-media platforms into censoring certain social-media content, in violation of the First Amendment.

*Missouri v. Biden*, slip. op. at 1, 83 F.4th 350 (5th Cir. 2023).

> [T]he platforms' policies were not clear-cut and did not always lead to content being demoted. So, the White House pressed the platforms.

*Id.*, slip. op. at 4.

> [T]he White House, acting in concert with the Surgeon General's office, likely (1) coerced the platforms to make their moderation decisions by way of intimidating messages and threats of adverse consequences, and (2) significantly encouraged the platforms' decisions by commandeering their decision-making processes, both in violation of the First Amendment.

*Id*., slip. op. at 42. The coerced moderation decisions constituted state action violating the First Amendment. *Id*., slip. op. at 60-61. "Social-media platforms' content-moderation decisions must be theirs and theirs alone." *Id*., slip. op. at 71.

The Supreme Court ruled the plaintiffs lacked standing to obtain relief against the social media platforms for the asserted conduct of the government concerning content moderation.

> To establish standing, the plaintiffs must demonstrate a substantial risk that, in the near future, they will suffer an injury that is traceable to a Government defendant and redressable by the injunction they seek. Because no plaintiff has carried that burden, none has standing to seek a preliminary injunction.

*Murthy v. Missouri*, slip op. at 1, 603 U.S. 43, 144 S. Ct. 1972 (2024),. The Supreme Court majority opinion concluded an Injunction against the Government defendants "is unlikely to affect the platforms' content-moderation decisions." *Id*., slip op. at 27.

Both the majority opinion and the dissent acknowledge the government cannot coerce censorship by social media platforms.

> Because we do not reach the merits, we express no view as to whether the Fifth Circuit correctly articulated the standard for when the Government transforms private conduct into state action.

*Id.*, slip op. at 7 n.3.

> The plaintiffs assert several injuries—their past social-media restrictions, current self-censorship, and likely social-media restrictions in the future. The requested judicial relief, meanwhile, is an injunction stopping certain Government agencies and employees from coercing or encouraging the platforms to suppress speech. *A court could prevent these Government defendants from interfering with the platforms' independent application of their policies*. But without evidence of continued pressure from the defendants, it appears that the platforms remain free to enforce, or not to enforce, those policies—even those tainted by initial governmental coercion.

*Id.*, slip op. at 26 (emphasis added). See *Id*. (Alito dissenting) slip. op. at 3 ("[G]overnment officials may not coerce private entities to suppress speech, see *National Rifle Association of America v. Vullo*, 602 U. S. 175 (2024), and that is what happened in this case.").

### D. Broad Immunity Was Perpetuated by *Stare Decisis*

The scope of Section 230 immunity is "robust". *Carafano v. Metrosplash.com, Inc.*, 339 F3d 1119, 1122, 1123 (9th Cir. 2003) (immunity for *invasion of privacy* and *right of publicity* claims). The online platform is generally immune from tort claims based on third party content. U.S. courts applying Section 230, consistently held defamation claims could not be brought against a platform based on social media posts by individual users. *E.g., Zeran v. America Online Inc*., 129 F.3d 327, 330-31 (4th Cir. 2004); *Rigsby v. GoDaddy, Inc*., 59 F.4th 998 (9th Cir. 2023) (Section 230 shields internet service provider from liability for "state-law claims for invasion of privacy, publicity, trade libel, libel, and violations of Arizona's Consumer Fraud Act"); *Austin v. Crystaltech Web Hosting*, 211 Ariz. 569, 574, 125 P.3d 389, 394 (2005) (defamation).

When the elements of Section 230 immunity are present, even cases asserting breach of contract are properly dismissed. *See, e.g., Brittain v. Twitter, Inc.*, 2019 WL 2423375 at *5 (N.D. Cal. June 10, 2019); *Yuksel v. Twitter Inc.*, 22-cv-05415-TSH, 5 (N.D. Cal. Nov. 7, 2022).

Some courts reached conflicting results on issues like immunity for claims alleging design defects on a social media platform. *Compare Lemmon v. Snap, Inc.*, 440 F. Supp. 3d 1103 (C.D. Cal. 2020), *rev'd,* 995 F.3d 1085, 1094 (9th Cir. 2021) (no immunity for claims arising out of "designing Snapchat in such a way that it allegedly encourages dangerous behavior"); *Maynard v. Snapchat, Inc.*, 346 Ga. App. 131, 816 S.E. 2d 77 (2018) (no immunity), *with Doe v. MySpace, Inc.*, 528 F.3d 413, 421 (5th Cir. 2008) (immunity for claimed "failure to implement measures that would have prevented Julie Doe from communicating with" her eventual attacker"); *Doe v. Snap, Inc.*, No. 22-20543, 2023 WL 4174061, slip op. at 2 (5th Cir. Jun. 26, 2023) (per curium, not designated for publication) (immunity for "claims under Texas law for negligent undertaking, negligent design, and gross negligence.").

> The *en banc* court, by a margin of one, voted against revisiting our erroneous interpretation of Section 230, leaving in place sweeping immunity for social media companies that the text cannot possibly bear.

*Doe v. Snap,* Inc., 88 F. 4th 1069, slip op. at 3 (5th Cir. 2023) (Elrod, C. J., dissenting from denial of rehearing *en banc*), cert. denied, S. Ct. Docket No. 23–961 (July 2, 2024) (Thomas, J., and Gorsuch, J. dissenting from the denial of certiorari) available at https://www.law.cornell.edu/supremecourt/text/23-961.

Judicial interpretation extended the scope of Section 230 immunity to cover recommending and targeting content related to terrorism. Facebook allegedly used algorithms that directed inflammatory Hamas postings to personalized newsfeeds of other users. Victims (or relatives of victims) claimed this function contributed to terrorist attacks. "Merely arranging and displaying others' content to users ... through such algorithms ... is not enough to hold [a service provider] responsible as the 'develop[er]' or 'creat[or]' of that content." *Force v. Facebook, Inc.,* 934 F.3d 53, 70 (2d Cir. 2019).

Google also allegedly used machine learning algorithms on its You Tube platform to provide access to terrorist propaganda. Plaintiffs suing on behalf of Noheme Gonzalez, deceased, claimed that Google aided and abetted an act of terrorism in Paris France. The Ninth Circuit found Google was entitled to Section 230 immunity. *Gonzalez v. Google LLC*, 2 F.4th 871, 892 (9th Cir. 2021), *vacated on other grounds and remanded*, 598 U.S. 617 (2023).

Once U.S. courts interpreted Section 230 broadly, *stare decisis* required adherence to that precedent.

> In sum, though we agree the Internet has grown into a sophisticated and powerful global engine the drafters of § 230 could not have foreseen, the decision we reach is dictated by the fact that we are not writing on a blank slate. Congress affirmatively immunized interactive computer service providers that publish the speech or content of others.

*Gonzalez v. Google LLC*, 2 F.4th at 896-97, *vacated on other grounds and remanded*, 598 U.S. 617 (2023).

> *Stare decisis* —in English, the idea that today's Court should stand by yesterday's decisions—is "a foundation stone of the rule of law." *Michigan v. Bay Mills Indian Community,* 572 U.S. [782], 134 S. Ct. 2024, 2036, 188 L.Ed.2d 1071 (2014)."

*Kimble v. Marvel Entm't, LLC*, 576 U.S. 446, 455, 135 S. Ct. 2401 (2015).

### E. Voluntary Removal of Defamatory Content Under Platform Terms of Use

Generally, the online platform will remove defamatory content under the terms of its user agreement. The protections of Section 230 continue to apply, however, "even if the website proprietor has not acted to remove offensive content posted by others." *Doe v. Internet Brands, Inc.*, 824 F.3d 846, 851 (9th Cir. 2016), *citing Barnes v. Yahoo!, Inc.*, 570 F.3d 1096, 1103 (9th Cir. 2009). *Accord Ricci v. Teamsters Union Local 456*, 781 F.3d 25, 28 (2d Cir. 2015) (no liability for merely refusing to remove an allegedly defamatory newsletter). The California Supreme Court concluded section 230 prevents ordering a social media platform to remove social media content even if a court has found the content defamatory and ordered it to be removed. *Hassell v. Bird*, 5 Cal.5th 522, 234 Cal. Rptr. 3d 867, 420 P.3d 77 (2018) (plurality opinion).

Some platforms actually stated a policy of not removing user content. For example,

> Xcentric operates a website called the Ripoff Report. The website's purpose is to permit consumers 'to post free complaints, called "reports," about companies and individuals whom [sic] they feel have wronged them in some manner.

*Small Justice LLC v. Xcentric Ventures LLC*, 873 F.3d 313, 316 (1st Cir. 2017). The Ripoff Report website Terms and Conditions stated:

> By posting this report/rebuttal, I attest this report is valid. I am giving Rip-off Report irrevocable rights to post it on the website. I acknowledge that once I post my report, it will not be removed, even at my request.

*Id*. Xcentric relied on Section 230 to justify this practice.

> Congress made a policy choice to protect online speech in enacting Section 230 and treating Xcentric as a content provider 'would flout Congress's intent by wrongly preventing … Xcentric from claiming immunity.'

*Id.*, at 322, *citing Universal Commc'n Sys., Inc. v. Lycos, Inc*., 478 F.3d 413, 418 (1st Cir. 2007).

### F. Reasons for Destination Law Not to Follow Section 230

There are at least two reasons Destination law might not apply broad Section 230 immunity. The U.S. Congress was unable to agree on amendments to Section 230 to permit liability for failure to moderate content. As a result, Internet content was regulated by the Internet platforms. Voluntary efforts by large ISPs to moderate content were perceived by some as unfair limitations on Internet user's free speech and by others as inadequate to protect against abusive and harmful content (such as defamation).

The "robust" immunity Section 230 provided to ISPs encouraged the growth of the Internet over decades as the technology advanced. Based on the unforeseen growth of social media technology, U.S. courts acknowledged Section 230 would benefit from further clarification by Congress or the Supreme Court. Until the law was changed or interpreted by higher authority, however, U.S. courts were bound by the established precedent to apply broad immunity to ISPs for injuries resulting to Internet users from Internet content. See **part I(C)(2)** above

> [S]ection 230 is no model of clarity, and there is ample room for disagreement about its scope. *See generally Malwarebytes, Inc. v. Enigma Software Grp. USA, LLC* , ––– U.S. ––––, 141 S. Ct. 13, 208 L.Ed.2d 197 (2020) (statement of Thomas, J., respecting denial of certiorari). Despite the statutory text's indeterminacy, the uniform view of federal courts interpreting this federal statute requires dismissal of claims alleging that interactive websites like Facebook should do more to protect their users from the malicious or objectionable activity of other users. The plaintiffs' claims for negligence, negligent undertaking, gross negligence, and products liability all fit this mold. The United States Supreme Court—or better yet, Congress—may soon resolve the burgeoning debate about whether the federal courts have thus far correctly interpreted section 230 to bar such claims. Nevertheless, the prevailing judicial interpretation of section 230 has become deeply imbedded in the expectations of those who operate and use interactive internet services like Facebook. We are not interpreting section 230 on a clean slate, and we will not put the Texas court system at odds with the overwhelming federal precedent supporting dismissal of the plaintiffs' common-law claims.

*In re Facebook, Inc.*, 625 S.W.3d 80, 83-84 (Tex. 2021).

> We share the dissent's concerns about the breadth of § 230. As the dissent observes, "there is a rising chorus of judicial voices cautioning against an overbroad reading of the

> scope of Section 230 immunity," and the feasibility of screening for dangerous content is being revisited. For example, websites are leveraging new technologies to detect, flag, and remove large volumes of criminal content such as child pornography. In light of the demonstrated ability to detect and isolate at least some dangerous content, Congress may well decide that more regulation is needed. In the meantime, our decision does not extend what the dissent rightly describes as § 230's sweeping scope.
>
> …
>
> In sum, though we agree the Internet has grown into a sophisticated and powerful global engine the drafters of § 230 could not have foreseen, the decision we reach is dictated by the fact that we are not writing on a blank slate. Congress affirmatively immunized interactive computer service providers that publish the speech or content of others.

*Gonzalez v. Google LLC*, 2 F.4th at 896-97, *vacated on other grounds and remanded*, 598 U.S. 617 (2023).

Destination does not face the constraints resulting from (1) legislation adopting the immunizing language of Section 230, or (2) prior case law interpreting Section 230 broadly thereby implicating the doctrine of *stare decisis*. We can write on a "clean slate."

### G. The Effect of Limiting Section 230 Immunity Under Destination Law

Although Section 230 was enacted specifically in the United States, and modified over time, core editorial decisions by ISPs on displaying or withdrawing user generated content remain protected. And to publish to the important U.S. market, ISPs relied on the protection of U.S. law. Critics disagree on whether the result is overprotection or inadequate protection of free speech and Internet user content

Destination law need not protect ISPs under an equivalent to Section 230, but almost all other jurisdictions do. See **Part I(C)** above regarding the European Union safe harbor regime.

Some U.S. courts have determined that Section 230 applies to conduct outside the U.S. when litigation is filed in U.S. courts.

> In other words, because § 230(c)(1) focuses on limiting liability, the relevant conduct occurs where immunity is imposed, which is where Congress intended the limitation of liability to have an effect, rather than the place where the claims principally arose. As such, the conduct relevant to § 230's focus is entirely within the United States—i.e., at the situs of this litigation.

*Gonzalez v. Google LLC*, 2 F.4th 871, 887 (9th Cir. 2021), *vacated on other grounds and remanded*, 598 U.S. 617 (2023) (application of Section 230 immunity to death from a terrorist attack in Paris

France—conduct entirely outside the U.S.—based on access to terrorist propaganda on Google). Changing Destination law will not prevent ISPs from filing suit and seeking immunity in the U.S.

Conversely, judgments from the Destination Court of Arbitration may not be enforced in the U.S. if there is no immunity for ISPs.

> Notwithstanding any other provision of Federal or State law, a domestic court shall not recognize or enforce a foreign judgment for defamation against the provider of an interactive computer service, as defined in section 230 of the Communications Act of 1934 (47 U.S.C. [§] 230 ) unless the domestic court determines that the judgment would be consistent with section 230 if the information that is the subject of such judgment had been provided in the United States.

28 U.S.C. § 4102(c)(1).

Thus, Section 230 applies to enforcement of a Destination Court of Arbitration Decision in the U.S., even if the conduct at issue occurred on Destination. Without Section 230 immunity, a Decision against an ISP can only be enforced on Destination (where the ISPs have limited assets).

### H. Anonymity Complicates Enforcement of Defamation Claims on Social Media

Anonymity features have long been accepted out of concern for internet privacy. *See, e.g., Dyroff v. Ultimate Software Grp., Inc.*, 934 F.3d 1093, 1100 (9th Cir. 2019) ("online privacy is a ubiquitous public concern for both users and technology companies"). Posting material anonymously "is an aspect of the freedom of speech protected by the First Amendment." *In re Anonymous Online Speakers*, 661 F.3d 1168, 1173 (9th Cir. 2011).

Before a suit can proceed, the victim must litigate whether the alleged wrongdoer is entitled to continued anonymity. *E.g., Best Western Int'l, Inc. v. Doe*, 2006 U.S. Dist. LEXIS 56014 (D. Ariz. July 25, 2006); *Mobilisa Inc. v. Doe*, 217 Ariz. 103, 170 P.3d 712 (App. 2007); *Columbia Ins. Co. v. seescandy.com*, 185 F.R.D. 573, 578-80 (N.D. Cal. 1999).

The individual user posting the content ultimately may be unable to satisfy a money judgment. And the same content can be reposted by other users and the process begins anew. This frustrating and ineffectual repetition was referred to as "Whack a Mole"—the name of a children's game chasing a mole as it popped up from various holes. https://idioms.thefreedictionary.com/whack-a-mole.

### Summary

To constitute defamation a false statement must "bring the defamed person into disrepute, contempt, or ridicule, or must impeach plaintiff's honesty, integrity, virtue, or reputation." This is a high bar. In addition, actual malice must be shown based on clear and convincing the speaker knew the statement was false or recklessly disregarded whether it might be false. *New York Times v. Sullivan*, 376 U.S. 254, 280, 84 S. Ct. 710 (1964). These issues will probably be disputed and result in arbitration. Moreover, the platform publishing the false statement will be immune from

suit and the individual responsible for the content will claim a right to anonymity. A suit for defamation is not a viable remedy to regulate objectionable social media content.

## II. THE COMMON LAW RIGHT OF PRIVACY

The U.S. Constitution does not mention privacy; consequently, the Supreme Court treated privacy rights as foundational assumptions in the shadow ("penumbra") of the Constitution itself. See *Griswold v. Connecticut*, 381 U.S. 479, 480, 85 S. Ct. 1678 (1965) (confirming the privacy of marital relations and use of contraceptives, the Court explained "specific guarantees in the Bill of Rights have penumbras, formed by emanations from those guarantees that help give them life and substance.").

Arizona is one of the first states whose founders adopted explicit protection for the privacy of its citizens. See Ariz. Const. art. 2, § 8. *Godbehere v. Phoenix Newspapers, Inc*., 162 Ariz. 335, 342, 783 P.3d 781, 788 (1989).

> [I]ndependently of the common rights of property, contract, reputation, and physical integrity, there is a legal right called the right of privacy ....

*Reed v. Real Detective Publ'g. Co*., 63 Ariz. 294, 305, 162 P.2d 133, 138 (1945).

Privacy protects against emotional, not reputational, harm. "The gravamen of [a privacy] action ... is the injury to the feelings of the plaintiff, the mental anguish and distress caused by the publication." *Reed*, 63 Ariz. at 305, 162 P.2d at 139.

> The remedy is available "to protect a person's interest in being left alone and is available when there has been publicity of a kind that is highly offensive."

*Godbehere*, 162 Ariz. at 341 (*quoting* Prosser & Keeton, On The Law Of Torts § 117 at 864 (5th ed. 1984)).

> the underlying objective of all four of the generally recognized invasion of privacy theories is to provide protection from 'interference with the interest of the individual in leading, to some reasonable extent, a secluded and private life, free from the prying eyes, ears and publications of others,' Restatement (Second) of Torts § 652A, comment b at 377 (1977), and to allow recovery for the mental distress occasioned by the interference. *See* Restatement § 652H.

*Rutledge v. Phoenix Newspapers, Inc*., 148 Ariz. 555, 558, 715 P.2d 1243 (App. 1986).

The common law right of privacy recognizes these four possible claims:

> Publication of private facts, *see, e.g., Reed v. Real Detective Publ'g. Co*., 63 Ariz. 294, 305, 162 P.2d 133, 138 (1945).

Intrusion upon seclusion, *see e.g., Hart v. Seven Resorts Inc*., 190 Ariz. 272, 279, 947 P.2d 846, 853 (App. 1997).

False light, *see e. g., Godbehere v. Phoenix Newspapers, Inc*., 162 Ariz. 335, 342, 783 P.3d 781, 788 (1989).

Right of publicity, see e.g., *In re Estate of Reynolds*, 235 Ariz. 80, 83 ¶12, 327 P.3d 213, 216 (App. 2014).

### A. Publication of Private Facts

"A person who unreasonably and seriously interferes with another's interest in not having his affairs known to others" might be held liable for invasion of privacy. *Reed v. Real Detective Pub. Co*., 63 Ariz. 294, 162 P.2d 133, 137 (1945).

> We have not previously considered the question of whether the requirements of intentional infliction of emotional distress should be extended to an action for invasion of privacy based on publicity given to private life. We conclude that such an extension is appropriate. This court first incorporated the requirements of the mental distress tort into an action for invasion of privacy based on intrusion of one's solitude or seclusion in *Cluff v. Farmers Ins. Exch*., 10 Ariz. App. 560, 460 P.2d 666 (App. 1969). We incorporated the safeguards of the tort into that particular invasion of privacy theory because the gist of the wrong redressed by that theory "is clearly the intentional infliction of mental distress." 10 Ariz. App. at 564, 460 P.2d at 670 quoting Prosser, Privacy, 48 Calif. L. Rev. 383, 422 (1960). The requirements, in other words, were incorporated into the intrusion theory of privacy to prevent that theory from being used as a means by which to circumvent the stringent standards necessary to otherwise establish a claim for the intentional infliction of emotional distress.

*Rutledge v. Phoenix Newspapers, Inc*., 148 Ariz. 555, 557, 715 P.2d 1243 (App. 1986).

### B. Intrusion Upon Seclusion

"One who intentionally intrudes, physically or otherwise, upon the solitude or seclusion of another or his private affairs or concerns, is subject to liability to the other for invasion of privacy, if the intrusion would be highly offensive to a reasonable person." *Hart v. Seven Resorts Inc*., 190 Ariz. 272, 279, 947 P.2d 846, 853 (App. 1997) (*quoting* Restatement (Second) of Torts § 652B).

Intrusion is not limited to physical entry.

> “It may also be by the use of the defendant's senses, with or without mechanical aids, to oversee or overhear the plaintiff's private affairs, as by looking into his upstairs window with binoculars or tapping his telephone wires. It may be by some other form of investigation or examination into his private concerns, as by opening his private and personal mail, searching his safe or wallet, examining his private bank account, or compelling him by a forged court order to permit an inspection of his personal documents.”

*Id.*, (quoting Restatement (Second) of Torts § 652B, comment b).

*Publication of private facts* and *intrusion upon seclusion* claims do not apply to the social media posts we are discussing. Both publication of private facts and intrusion claims require outrageous conduct.

> Arizona cases applying the intentional infliction "extreme and outrageous conduct" standard to invasion of privacy claims include: ... *Rutledge v. Phoenix Newspapers, Inc*., 148 Ariz. 555, 715 P.2d 1243 (App. 1986) (private facts, false light); *Creamer v. Raffety*, 145 Ariz. 34, 699 P.2d 908 (App. 1985) (intrusion)

*Godbehere*, 162 Ariz. at 339 n.1, 83 P.2d at 783.

C. **False light invasion of privacy.**

> One who gives publicity to a matter concerning another that places the other before the public in a false light is subject to liability to the other for invasion of his privacy, if
>
> (a) the false light in which the other was placed would be highly offensive to a reasonable person, and
>
> (b) the actor had knowledge of or acted in reckless disregard as to the falsity of the publicized matter and the false light in which the other would be placed.

Restatement (Second) of Torts § 652E. *See Godbehere,* 162 Ariz. at 338, 783 P.2d at 784.

Proof of false light, however, does not require a false statement.

> Although a cause of action for false light invasion of privacy may arise when someone publishes something untrue about a person, in some instances, even a true statement may form the basis for false light liability if it creates a false implication about the person. See *Godbehere*, 162 Ariz. at 341, 783 P.2d at 787 ("[T]he false innuendo created by the highly offensive presentation of a true fact constitutes the injury." (citing Restatement (Second) of Torts § 652E)).

*Desert Palm Surgical Grp., PLC v. Petta*, 236 Ariz 568, 580, 343 P.3d 438, 450 (App. 2015).

**Actual Malice:** A false light invasion of privacy claim requires proof of publication by another (1) with actual malice, (2) placing you "in a false light highly offensive to a reasonable person," including (3) "a major misrepresentation of [your] character, history, activity, or beliefs." *Canas v. Bay Entm't*, 252 Ariz. 117, 498 P.3d 1082, 1087 (App. 2021).

> Under this theory, a plaintiff may recover even in the absence of reputational damage, as long as the publicity is unreasonably offensive and attributes false characteristics. However, to qualify as a false light invasion of privacy, the publication must involve "a major

> misrepresentation of [the plaintiff's] character, history, activities or beliefs," not merely minor or unimportant inaccuracies. Restatement (Second) of Torts § 652E comment c.

*Godbehere*, 162 Ariz. at 341.

> Under false light, "public" means the matter is communicated to "the public at large" or "to so many persons that the matter must be regarded as substantially certain to become one of public knowledge." Restatement (Second) of Torts §§ 652D cmt. a, 652E cmt. a. Appellant has not alleged that the purportedly actionable statements were made to the public at large, that they would be highly offensive to a reasonable person, or that Appellees acted intentionally or with reckless disregard in making the statements. The superior court properly dismissed the invasion of privacy claims

*Snyder v. Banner Health, an Ariz. Corp.*, No. 1 CA-CV 13-0630, 2014 WL 4980382 (Ariz. Ct. App. Oct. 7, 2014) (Memorandum Decision)

**Public Officials:** As public officials, you cannot sue based on social media content related to your public duties or public life.

> Certainly, the public has a legitimate interest in the manner in which law enforcement officers perform their duties. Therefore, we hold that there can be no false light invasion of privacy action for matters involving official acts or duties of public officers.
>
> Consequently, we adopt the following legal standard: *a plaintiff cannot sue for false light invasion of privacy if he or she is a public official, and the publication relates to performance of his or her public life or duties.* We do not go so far as to say, however, that a public official has no privacy rights at all and may never bring an action for invasion of privacy. Certainly, if the publication presents the public official's *private life* in a false light, he or she can sue under the false light tort, although actual malice must be shown.

*Godbehere*, 162 Ariz. at 344 (emphasis added).

### D. The Right of Publicity

The right of publicity protects commercial interests in fame represented by name, likeness and image and is treated as a property right originating in unfair competition, intellectual property, and tort law. J. Thomas McCarthy, The Rights of Publicity and Privacy 1.7 (2d ed. 2013); *accord In Re Estate of Reynolds*, 235 Ariz. 80, 83 ¶15, 327 P.3d 213 (App. 2014)

The "actual malice" standard does *not* apply to the tort of appropriation of a right of publicity. *Zacchini* v. *Scripps-Howard Broadcasting Co.*, 433 U.S. 562, 97 S. Ct. 2849 (1977). The right of publicity protects the "commercial value of a person's identity" by prohibiting unauthorized use of "the person's name, likeness, or other indicia of identity for purposes of trade." Restatement (Third) of Unfair Competition §46 (1995). The right of publicity prevents others from diluting the value, or appropriating of the value, of a person's identity. *Id.* comment c. Proof of deception or confusion is not required. *Id.*

The social media publications use photographs (an image or likeness) of you both as well as your names. This is a misappropriation of the right of publicity. *See e.g., Canas v. Bay Entertainment, LLC*, 252 Ariz. 117, 498 P.3d 1082 (App. 2021) (photos of models used on social media to advertise nightclub).

**Commercial Use**. The commercial requirement use is missing. Typically, commercial use, or use “for purposes of trade,” is use for “advertising the user’s goods or services.” Restatement (Third) of Unfair Competition §47. “Use ‘for purposes of trade’ does not ordinarily include the use of a person’s identity in news reporting, commentary, entertainment, works of fiction or nonfiction, or in advertising that is incidental to such uses.” *Id.* For example, a sensationalized biography of Marilyn Monroe was not actionable. *Frosch v. Grosset & Dunlap, Inc*., 75 A.D.2d 768, 768-69 (N.Y App. Div. 1980).

### E. Application of Section 230 Immunity

Section 230 provides immunity from a lawsuit when liability would be based on publishing third party content. The scope of Section 230 is “robust”. *Rigsby v. GoDaddy, Inc*., 59 F.4th 998 (9th Cir. 2023) (Section 230 shields internet service provider from liability for “state-law claims for invasion of privacy, publicity, trade libel, libel, and violations of Arizona’s Consumer Fraud Act”); *Carafano v. Metrosplash.com, Inc.*, 339 F.3d 1119, 1122, 1123 (9th Cir. 2003) (immunity for invasion of privacy and right of publicity claims).

## Summary

A *false light* claim requires outrageous conduct and actual malice – the same high evidentiary burdens of a defamation claim. Moreover, at least one case has held that *a plaintiff cannot sue for false light invasion of privacy if he or she is a public official and the publication relates to performance of his or her public life or duties. Godbehere v. Phoenix Newspapers, Inc*., 162 Ariz. 335, 344, 783 P.3d 781, 790 (1989). Unless the material relates to your personal life, a false light claim is not a viable remedy to regulate objectionable social media content. Even a claim regarding your private life is likely to be arbitrated due to disputed issues of material fact concerning outrageous conduct and actual malice.

The *right of publicity* protects commercial interests in use of your name, image, or likeness for “purposes of trade”. No actual malice is required, but “news reporting, commentary, entertainment, works of fiction or nonfiction” are not used for purposes of trade. Use of your images to advertise a revenue generating website would present the best claim, but the defendant will still try to create a disputed issue of material fact by claiming use for “news reporting, commentary, or entertainment.” The result probably would be arbitration.

---