

**GENERATIVE AI WILL BREAK THE INTERNET:  
BEYOND SECTION 230**

Graham H. Ryan\*

TABLE OF CONTENTS

I. INTRODUCTION .....	1
II. SECTION 230 GENERALLY .....	3
A. <i>What is an “interactive computer service”?</i> .....	4
B. <i>What is a “publisher or speaker”?</i> .....	5
C. <i>What is “another information content provider”?</i> .....	6
III. GENERATIVE AI .....	6
IV. A BREAKING POINT: TWO SECTION 230 TRENDS .....	10
A. <i>Whether Generative AI Systems are “Information Content Providers” under Section 230</i> .....	10
B. <i>Whether Generative AI Systems Fall Outside the Scope of Section 230 “Publisher or Speaker” Protection</i> .....	15
V. CONCLUSION .....	19

I. INTRODUCTION

The law that “created the Internet”<sup>1</sup> has reached a breaking point. Section 230 of the Communications Decency Act is a law enacted in 1996 that catalyzed the Internet’s development by providing Internet-based service providers like search engines and social networks legal immunity from lawsuits based on harmful content created by third parties.<sup>2</sup> For example, Google generally cannot be held civilly liable for simply retrieving and displaying harmful Internet-based content that it did not create.<sup>3</sup> Similarly, Facebook can typically avoid liability for harmful content contained in a user’s post on its platform.<sup>4</sup>

---

\* Graham H. Ryan, J.D., is a litigation and appellate partner at Jones Walker LLP, with extensive experience litigating complex commercial issues in all phases of litigation, appeals, and regulatory proceedings, including those involving technology, artificial intelligence, Section 230, and related matters. He holds an international designation as an Artificial Intelligence Governance Professional from the International Association of Privacy Professionals, and has been published on legal and constitutional issues arising from emerging Internet technologies.

1. See JEFF KOSSEFF, *THE TWENTY-SIX WORDS THAT CREATED THE INTERNET* (2019).  
2. 47 U.S.C. § 230 (c)(1) (generally, “Section 230”).  
3. See *Marshall’s Locksmith Serv. v. Google, LLC*, 925 F.3d 1263 (D.C. Cir. 2019).  
4. See *Force v. Facebook, Inc.*, 934 F.3d 53 (2d Cir. 2019).

For decades, U.S. courts have applied Section 230 protection expansively, furthering the law’s purpose “to preserve the vibrant and competitive free market that presently exists for the Internet and other interactive computer services, unfettered by Federal or State regulation.”<sup>5</sup> The functional scope of Section 230 immunity has evolved alongside Internet technology. It initially covered passive intermediaries like AOL online message boards at a time when the number of Internet users totaled 40 million,<sup>6</sup> but it now protects advanced social media algorithms that filter, promote, and personalize content<sup>7</sup> as the population of Internet users has surpassed 5.35 billion.<sup>8</sup> In the intervening period, courts have stretched Section 230 to its logical bounds — and some would argue far beyond.<sup>9</sup> But the law establishes one bright line courts have not crossed: Section 230 protection does not extend to Internet-based services that actually create or develop content.<sup>10</sup>

Technology has crossed that line. The Internet’s future development will be shaped by generative artificial intelligence (AI), which performs an unprecedented technological role in creating and developing content rather than merely retrieving or exchanging it.<sup>11</sup> Courts have long been reluctant to disrupt Section 230’s legal underpinnings of the Internet<sup>12</sup> and have carefully adapted Section 230’s legal standards to extend its protections to new

---

5. 47 U.S.C. § 230(b)(2).

6. *Zeran v. Am. Online, Inc.*, 129 F.3d 327, 328 (4th Cir. 1997) (“‘The Internet is an international network of interconnected computers,’ currently used by approximately 40 million people worldwide.”) (citing *Reno v. Am. C.L. Union*, 521 U.S. 844, 849 (1997)).

7. *Force*, 934 F.3d 53.

8. See Ani Petrosyan, *Number of Internet and Social Media Users Worldwide as of January 2024*, STATISTA (Jan. 31, 2024), <https://www.statista.com/statistics/617136/digital-population-worldwide/> (“As of January 2024, there were 5.35 billion internet users worldwide, which amounted to 66.2 percent of the global population. Of this total, 5.04 billion, or 62.3 percent of the world’s population, were social media users.”).

9. *Doe v. Snap, Inc.*, No. 22-20543, 2023 U.S. App. LEXIS 33501, at \*2 (5th Cir. Dec. 18, 2023) (Elrod, J., dissenting) (arguing that the current scope of Section 230 protection provides “sweeping immunity for social media companies that the text [of Section 230] cannot possibly bear”).

10. 47 U.S.C. § 230(c)(1) (conferring immunity only where “information [is] provided by *another* information content provider.” (emphasis added)); 47 U.S.C. § 230(f)(3) (defining “information content provider” to include “any person or entity that is responsible, in whole or in part, for the *creation or development of information*” (emphasis added)).

11. See *infra* Section III; see, e.g., *TAMBIAMA MADIEGA*, EUR. PARL. RSCH. SERV., NO. PE 745.708, GENERAL-PURPOSE ARTIFICIAL INTELLIGENCE (2023).

12. See, e.g., Transcript of Oral Argument at 54, 82, *Gonzalez v. Google LLC*, 598 U.S. 617 (2023) (No. 21-1333) (Kavanaugh, J.) (“Congress drafted a broad text, and that text has been unanimously read by courts of appeals over the years . . . [I]sn’t it better for — to keep it the way it is, for us, and Congress — to put the burden on Congress to change that and they can consider the implications and make these predictive judgments?”); *id.* at 46 (Kagan, J.) (“[I]sn’t that something for Congress to do, not the Court?”).

technologies.<sup>13</sup> But courts will soon be confronted with unavoidable calls to reshape Section 230 in the context of whether it applies to protect the generative AI systems that will drive the Internet’s future.

Litigation involving generative AI will force courts to revisit Section 230 in a manner that will vastly impact the Internet in two specific ways. First, court opinions on generative AI are likely to expressly declare that Section 230’s legal protections do not extend to AI systems that materially contribute to the creation or development of content, thereby increasing legal risk and thwarting growth of Internet-based generative AI systems.<sup>14</sup> Second, court opinions on generative AI are likely to reshape two legal standards that govern the scope of Section 230 immunity — the meaning of an “information content provider,” and the contours of when a technology service is a “publisher or speaker” of information. This will disrupt longstanding legal precedent and introduce legal risk for not only generative AI systems, but virtually all interactive computer services.<sup>15</sup>

Judicial determinations on Section 230 and generative AI will forever transform the legal landscape governing Internet technologies and dictate whether the legal protections that catalyzed the Internet’s pre-AI development will do the same for emerging technologies like generative AI. This article provides a primer on Section 230 and generative AI, and examines two evolving Section 230 legal standards that will soon be applied to generative AI in manner that will shape the future of the Internet.

## II. SECTION 230 GENERALLY

Section 230 states: “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.”<sup>16</sup> It provides that “no liability may be imposed under any State or local law” if a legal claim seeks to hold a provider liable for information created by a third-party.<sup>17</sup> For example, Section 230 has provided immunity to Facebook in a claim that it unlawfully provided Hamas, a U.S.-designated foreign terrorist organization, with a communications platform that enabled certain terrorist attacks committed by

---

13. *Force v. Facebook, Inc.*, 934 F.3d 53, 67–68 (2d Cir. 2019) (extending Section 230 immunity to claims based on algorithms that personalized Facebook content).

14. *See infra* Part IV.

15. *See id.*

16. 47 U.S.C. § 230(c)(1).

17. 47 U.S.C. § 230(e) (delineating exceptions to Section 230 immunity for federal criminal prosecution or claims under certain intellectual property law, electronic communications privacy law, or sex trafficking law).

Hamas in Israel.<sup>18</sup> Other examples include immunity for a website that allegedly facilitated illegal drug sales where the website did not create the content,<sup>19</sup> and immunity for a search engine in a claim based on its failure to remove an application from its app store.<sup>20</sup> Without Section 230, Internet-based services would be exposed to claims merely for displaying, transmitting, or blocking content created by third parties.<sup>21</sup> The Internet’s “uninhibited, robust” technological development is due in large part to the protections conferred by Section 230.<sup>22</sup> Section 230 “made e-commerce itself economically feasible”<sup>23</sup> and expansively catalyzed the technological development of the Internet.<sup>24</sup>

Section 230 immunity generally applies when three criteria are met: (1) the provider is an “interactive computer service,” (2) a claim treats the provider as the “publisher or speaker” of harmful information, and (3) the harmful information is created by “another information content provider.”<sup>25</sup>

#### A. What is an “interactive computer service”?

Section 230 broadly defines “interactive computer service” to include “any information service, system, or access software provider that provides or enables computer access by multiple users to a computer server.”<sup>26</sup> Courts

18. *Force*, 934 F.3d at 65–68.

19. *Dyroff v. Ultimate Software Grp., Inc.*, 934 F.3d 1093 (9th Cir. 2019).

20. *Ginsberg v. Google Inc.*, 586 F. Supp. 3d 998 (N.D. Cal. 2022).

21. *See, e.g., Google LLC v. Equustek Sols. Inc.*, No. 17-04207, 2017 U.S. Dist. LEXIS 182194, at \*6–7 (N.D. Cal. Nov. 2, 2017) (“The Ninth Circuit has held that, regardless of the underlying cause of action, a claim treats an intermediary as a publisher when it requires the intermediary to remove third-party content.”).

22. *See Jones v. Dirty World Entm’t Recordings LLC*, 755 F.3d 398, 415 (6th Cir. 2014) (citing 47 U.S.C. § 230(a)(1)–(5)) (“Congress envisioned an uninhibited, robust, and wide-open internet.”).

23. *Papataros v. Amazon.com, Inc.*, No. 17-9836, 2019 U.S. Dist. LEXIS 144253, at \*46 n.18 (D.N.J. Aug. 26, 2019) (“It has been said that the ‘twenty-six words’ of Section 230 of the CDA, enacted in 1996, made e-commerce itself economically feasible by permitting platforms such as Amazon.com to match sellers with buyers without taking on the seller’s liabilities. It would perhaps be more sober and accurate to say that the twenty-six words of Section 230 promoted or facilitated important aspects of the internet as we now know it.” (citation omitted)).

24. *See Doe v. MySpace Inc.*, 528 F.3d 413, 418 (5th Cir. 2008) (“Courts have construed the immunity provisions in § 230 broadly in all cases arising from the publication of user-generated content.”); *see also Diez v. Google, Inc.*, 831 F. App’x 723, 725 (5th Cir. 2020) (“A majority of federal circuits have interpreted § 230 ‘federal immunity’ to be rather broad.” (citing *Almeida v. Amazon.com, Inc.*, 456 F.3d 1316, 1321 (11th Cir. 2006))).

25. 47 U.S.C. § 230(c)(1); *Rigsby v. GoDaddy Inc.*, 59 F.4th 998, 1003 (9th Cir. 2023).

26. 47 U.S.C. § 230(f)(2); *see also* 47 U.S.C. § 230(f)(4) (“The term ‘access software provider’ means a provider of software (including client or server software), or enabling tools that do any one or more of the following: (A) filter, screen, allow, or disallow content; (B) pick, choose, analyze, or digest content; or (C) transmit, receive, display, forward, cache, search, subset, organize, reorganize, or translate content.”).

have applied this definition “expansively” to include social media platforms, search engines, online communities, and a wide range of other intermediary platforms and services that allow for the exchange and transmission of information between users.<sup>27</sup> Many generative AI systems, particularly those that employ Internet-based datasets, likely fall within Section 230’s definition of an interactive computer service.<sup>28</sup>

*B. What is a “publisher or speaker”?*

Section 230 immunizes a provider of an interactive computer service from a claim that attempts to hold it liable as the “publisher or speaker” of information created or developed by another person.<sup>29</sup> Section 230 does not define “publisher or speaker,” but court opinions generally hold that a legal claim impermissibly attempts to hold a provider liable as a publisher or speaker if the claim attempts to fault the provider for the third-party information,<sup>30</sup> would require the provider to edit or remove third-party content,<sup>31</sup> or is based on the provider’s failure to implement measures to prevent transmission of third-party content.<sup>32</sup> As one court explained, a claim attempts to hold a defendant liable as a “publisher or speaker” if the claim attempts to impose a duty on defendant to “police its network for content transmitted by its users.”<sup>33</sup> As another court explained, “any activity that can

---

27. *E.g.*, *Kimzey v. Yelp! Inc.*, 836 F.3d 1263, 1268 (9th Cir. 2016) (citing *Carafano v. Metroplash.com, Inc.*, 339 F.3d 1119, 1123 (9th Cir. 2003)) (“Yelp is plainly a provider of an ‘interactive computer service,’ a term that we interpret ‘expansive[ly]’ under the CDA.”). “[T]he most common interactive computer services are websites.” *Id.* (quoting *Fair Hous. Council v. Roommates.com, LLC*, 521 F.3d 1157, 1162 n.6 (9th Cir. 2008)).

28. *Id.*

29. 47 U.S.C. § 230(c)(1).

30. *See Doe v. Snap, Inc.*, No. 22-00590, 2022 U.S. Dist. LEXIS 119560, at \*43 (S.D. Tex. July 7, 2022) (seeking to hold defendant liable “as the publisher or speaker” of information “provided by a third party” (cleaned up)).

31. *See Bride v. Snap Inc.*, No. 21-06680, 2023 U.S. Dist. LEXIS 5481, at \*19 (C.D. Cal. Jan. 10, 2023) (“[T]he court finds that Plaintiffs’ theory would require the editing of third-party content, thus treating Defendants as a publisher of content.”); *Google LLC v. Equustek Sols. Inc.*, No. 17-04207, 2017 U.S. Dist. LEXIS 182194, at \*6–7 (N.D. Cal. Nov. 2, 2017) (“The Ninth Circuit has held that, regardless of the underlying cause of action, a claim treats an intermediary as a publisher when it requires the intermediary to remove third-party content.”).

32. *See Doe v. MySpace Inc.*, 528 F.3d 413, 420 (5th Cir. 2008) (“Their claims are barred by the CDA, notwithstanding their assertion that they only seek to hold MySpace liable for its failure to implement measures that would have prevented [the communication]. Their allegations are merely another way of claiming that MySpace was liable for publishing the communications and they speak to MySpace’s role as a publisher of online third-party-generated content.”).

33. *Green v. Am. Online, Inc.*, 318 F.3d 465, 470–71 (3d Cir. 2003) (finding that Section 230 barred claim where the plaintiff alleged that the defendant had negligently failed “to properly police its network for content transmitted by its users” because the plaintiff “attempt[ed] to hold [the defendant]

be boiled down to deciding whether to exclude material that third parties seek to post online is perforce immune under section 230.”<sup>34</sup>

### C. What is “another information content provider”?

Section 230 immunity applies when harmful information was created or developed by a third-party “information content provider,” defined as “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.”<sup>35</sup> Section 230 immunity thus does not apply if a provider of an interactive computer service itself is the “information content provider,” i.e., if the provider materially contributes to the creation or development of the harmful content.<sup>36</sup>

## III. GENERATIVE AI

The definitions for AI and its constituent terms are varying and evolving, and there is “no globally agreed definition of artificial intelligence.”<sup>37</sup> Generally, the term “artificial intelligence” refers to computer-based systems that use machine and human inputs to perceive real and virtual environments, abstract perceptions into models through automated analysis, and use model inference to formulate options.<sup>38</sup> An “AI model” refers to a component of an

---

liable for decisions relating to the monitoring, screening, and deletion of content from its network—actions quintessentially related to a publisher’s role.”)

34. *Fair Hous. Council v. Roommates.com, LLC*, 521 F.3d 1157, 1170–71 (9th Cir. 2008); *see also Doe v. Twitter, Inc.*, Nos. 22-15103, 22-15104, 2023 U.S. App. LEXIS 10808, at \*4 (9th Cir. May 3, 2023).

35. 47 U.S.C. § 230(c)(1); *see also* 47 U.S.C. § 230(f)(3).

36. 47 U.S.C. § 230(c)(1) (conferring immunity only where “information [is] provided by *another* information content provider.”) (emphasis added); *see also Roommates*, 521 F.3d at 1166 (“[S]ection 230 provides immunity only if the interactive computer service does not ‘creat[e] or develop[]’ the information ‘‘in whole or in part.’” (citing 47 U.S.C. § 230(f)(3))); *Kimzey v. Yelp! Inc.*, 836 F.3d 1263, 1269 (9th Cir. 2016) (“[A] website may lose immunity under the CDA by making a material contribution to the creation or development of content.”).

37. *See MADIEGA*, *supra* note 11, at 1 (“[T]here is no globally agreed definition of artificial intelligence.”).

38. *See* 15 U.S.C. § 9401(3); National Artificial Intelligence Initiative Act of 2020, Pub. L. No. 116-283, 134 Stat. 3388; *see also* Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules On Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, EUR. PARL. DOC. 2021/0106 (COD), art. 3 (“[A] system that is designed to operate with elements of autonomy and that, based on machine and/or human-provided data and inputs, infers how to achieve a given set of objectives using machine learning and/or logic- and knowledge based approaches, and produces system-generated outputs such as content (generative AI systems), predictions, recommendations or decisions, influencing the environments with which the AI system interacts.”); Digital Charter Implementation Act, S.C. 2022, c C-27 (Can.) (“A technological system that, autonomously or partly autonomously, processes data related to human activities through

information system that produces outputs from a set of inputs through AI technology and computational or machine-learning techniques.<sup>39</sup> “Machine learning” is an application of AI that provides a system the ability to automatically improve from data or experience, without explicit programming.<sup>40</sup> An “AI system” generally refers a system, application, or tool that operates using AI.<sup>41</sup>

“Generative AI” generally refers to a subset of AI models that generate content derived through machine learning, input data, and pre-existing data.<sup>42</sup> Many generative AI systems employ large language models (“LLMs”) that are trained on a large dataset of text from the Internet to predict the next plausible word or phrase in a linguistic construct.<sup>43</sup> Some LLMs are fine-tuned using additional data and a technique called reinforcement learning from human feedback (“RLHF”), which uses human augmentation to guide model outputs.<sup>44</sup> Model tuning and training can continue after AI system deployment to an end user through filtering tools and content moderation classifiers.<sup>45</sup> Foundation models employed by generative AI systems usually generate original outputs, as contrasted from systems like search engines that may retrieve and display extracted third-party content from a source website. While the particulars of the definitions and intricacies of generative AI vary, one material aspect of generative AI for Section 230 purposes is that the content output from some generative AI systems, which can include images,

---

the use of a genetic algorithm, a neural network, machine learning or another technique in order to generate content or make decisions, recommendations or predictions.”).

39. *See, e.g.*, Exec. Order No. 14110, 88 Fed. Reg. 75191, 75193 (Oct. 30, 2023) [hereinafter AI Order] (Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence).

40. *See, e.g.*, 15 U.S.C. § 9401(11).

41. *See, e.g.*, AI Order, *supra* note 39, at 75193.

42. *See, e.g., id.* at 75195; *see also Key Terms for AI Governance*, INT’L ASS’N PRIV. PROS. (Nov. 2023), <https://iapp.org/resources/article/key-terms-for-ai-governance/> (“Generative AI” is “[a] field of AI that uses deep learning trained on large datasets to create new content, such as written text, code, images, music, simulations and videos . . . [that] makes predictions on existing data rather than new data [and is] capable of generating novel outputs based on input data or user prompts.”). *See also Generative AI Overview*, GOOGLE AI, <https://ai.google/discover/generativeai> (last visited Mar. 23, 2024) (“Generative AI builds on existing technologies, like large language models (LLMs) which are trained on large amounts of text and learn to predict the next word in a sentence. For example, ‘peanut butter and \_\_\_’ is more likely to be followed by ‘jelly’ than ‘shoelace’. Generative AI can not only create new text, but also images, videos, or audio.”).

43. OPENAI, GPT-4 SYSTEM CARD 2 (2023).

44. *Id.*; *see, e.g.*, Long Ouyang et al., *Training Language Models to Follow Instructions with Human Feedback*, 35 ADVANCES NEURAL INFO. PROCESSING SYS. 27730 (2022); Paul Roit et al., *Factually Consistent Summarization via Reinforcement Learning with Textual Entailment Feedback*, 61 ASS’N FOR COMPUTATIONAL LINGUISTICS 6252 (2023).

45. OPENAI, *supra* note 43, at 26.

videos, audio, text, and other digital content, may be contributed to, or developed in part by the AI system itself.<sup>46</sup>

Internet technologies are rapidly embracing generative AI. Research laboratory OpenAI has released various AI models that have quickly evolved in technological advancement.<sup>47</sup> In 2020, OpenAI released GPT-3, a language model trained on Internet data that can perform various natural language processing tasks, such as summarization and translation.<sup>48</sup> In 2021, OpenAI released DALL-E, a deep-learning model that can generate digital images from natural language prompts. In December 2022, OpenAI released ChatGPT, a chatbot that can generate text using Internet data and machine learning models. In 2023, OpenAI launched GPT-4, a general-purpose AI tool with a wide range of interdisciplinary functions.<sup>49</sup> In 2024, OpenAI debuted Sora, an AI model that can create realistic and imaginative scenes from text instructions.<sup>50</sup> Search engines, web browsers, and other interactive computer services are increasingly incorporating generative AI and making AI tools widely available to the public.<sup>51</sup>

Societal understanding of the risks and harms posed by AI systems is emerging alongside AI technology.<sup>52</sup> Generative AI presents the “potential for vast societal impacts,” and AI model developers admit that AI limitations and capabilities present “safety challenges.”<sup>53</sup> Risks and potential harms identified by generative AI developers include: hallucinations (i.e., “content that is nonsensical or untruthful in relation to certain sources”);<sup>54</sup> harmful content (i.e., “content that may pose harm to individuals, groups, or society” such as “hate speech, discriminatory language, incitements to violence, or content that is then used to either spread false narratives or to exploit an individual”);<sup>55</sup>

---

46. *Id.*; *cf.* Thaler v. Vidal, 43 F.4th 1207, 1213 (Fed. Cir. 2022) (holding that “only a natural person can be an inventor, so AI cannot be” and stating that “we are not confronted today with the question of whether inventions made by human beings with the *assistance* of AI are eligible for patent protection” (emphasis original)).

47. See MADIEGA, *supra* note 11, at 1.

48. *Id.*

49. *Id.*

50. Sora, OPENAI, <https://openai.com/sora> (last visited Mar. 23, 2024).

51. See MADIEGA, *supra* note 11, at 1 (discussing Microsoft and Google each launching AI-powered search tools); see also GOOGLE AI, <https://ai.google/> (last visited Mar. 23, 2024).

52. See INT’L ASS’N PRIV. PROS., AIGP BODY OF KNOWLEDGE 7 (2023) (outlining potential risks and harms arising from AI, including harm to individuals (*e.g.*, issues relative to privacy, civil rights, economic opportunity, and bias in employment, housing, education, or pricing); harm to groups (*e.g.*, discrimination toward sub-groups, mass surveillance, tracking, and profiling); harm to society (*e.g.*, issues undermining the democratic process, public trust in governmental institutions, educational access, and job redistribution); harm to a organizations and institutions (*e.g.*, reputational, cultural, and economic damage); and harm to ecosystems (*e.g.*, impact on supply chains and natural resources)).

53. OPENAI, *supra* note 43, at 1.

54. *Id.* at 6.

55. *Id.* at 7.



disinformation and influence operations (e.g., “generating content that is intended to mislead” including “plausibly realistic and targeted content, including news articles, tweets, dialogue, and emails”);<sup>56</sup> proliferation of conventional and unconventional weapons (e.g., “suggest[ing] vulnerable public targets” and “generat[ing] the fundamental components that are required to engineer a radiological dispersal device”);<sup>57</sup> privacy (e.g., model training on “licensed, created, and publicly available data sources, which may include publicly available personal information”);<sup>58</sup> cybersecurity (i.e. “social engineering (like drafting phishing emails), and explaining some vulnerabilities”);<sup>59</sup> economic impacts (e.g., “models may lead to the automation of certain jobs” that “could result in workforce displacement”);<sup>60</sup> harms of representation (i.e., perpetuation, reinforcement, or application of social biases, stereotypes, and worldviews contained in a dataset or prompt);<sup>61</sup> allocation (i.e., allocation of opportunities or resources in a manner based on social biases, stereotypes, and worldviews);<sup>62</sup> acceleration (i.e., “the risk of [AI development] racing dynamics leading to a decline in safety standards, the diffusion of bad norms, and accelerated AI timelines, each of which heighten societal risks associated with AI”);<sup>63</sup> and overreliance (i.e., “when users excessively trust and depend on the model, potentially leading to unnoticed mistakes and inadequate oversight.”)<sup>64</sup> AI developers have identified some technical mitigation strategies, including adversarial testing of models through “red teaming” to determine whether models are likely to produce harmful content;<sup>65</sup> reducing the prevalence of certain types of content in the pre-training dataset;<sup>66</sup> fine-tuning a model to refuse certain instructions such as direct requests for illicit advice;<sup>67</sup> leveraging data from inputs to anticipate and reduce adversarial prompting;<sup>68</sup> and enforcing use policies

---

56. *Id.* at 10. AI systems produce content outputs that are increasingly “more believable and more persuasive.” *Id.* at 4.

57. *Id.* at 12.

58. *Id.* at 13.

59. *Id.* “[AI model development] does continue the trend of potentially lowering the cost of certain steps of a successful cyberattack, such as through social engineering or by enhancing existing security tools.” *Id.* at 3.

60. *Id.* at 18

61. *Id.* at 7.

62. *Id.* at 9. Generative AI “can rival human propagandists in many domains, especially if teamed with a human editor” *Id.* at 10. AI systems produce content outputs that are increasingly “more believable and more persuasive.” *Id.* at 4.

63. *Id.* at 19.

64. *Id.*

65. *Id.* at 4–5.

66. *Id.* at 3.

67. *Id.*

68. *Id.*

through training and monitoring for new risks.<sup>69</sup> AI system-level mitigations can also be employed through AI governance strategies, implementing best practices and policies, and monitoring for policy violations.<sup>70</sup>

#### IV. A BREAKING POINT: TWO SECTION 230 TRENDS

Two Section 230 issues will shape legal risk and its resulting impact on the growth of generative AI and Internet technologies. The first issue is how the legal standards for the “information content provider” test will apply to generative AI systems that contribute to the creation or development of content. The second issue is how the legal standards for the “publisher or speaker” test will apply to the novel and evolving design and functionality of generative AI systems, particularly whether Section 230 protects generative AI systems from claims based on negligent design and product liability.

Judicial analysis of these issues may turn on several technical yet evolving criteria related to the design, development, and deployment of a particular AI system. This analysis touches on training data, model type, model tuning, the system’s use of algorithms and machine learning, the role of user prompts and inputs, and the level and nature of contribution of a system’s various components to the creation or development of the content giving rise to a particular claim, among other factors discussed below.

##### *A. Whether Generative AI Systems are “Information Content Providers” under Section 230*

Section 230 immunity does not extend to an interactive computer service that is “responsible, in whole or in part, for the creation or development” of harmful information giving rise to a claim.<sup>71</sup> Court opinions provide helpful but evolving guidelines for applying this statutory text. Courts generally apply the “material contribution” test, which examines the extent to which a provider contributed to the creation or development of the information at issue.<sup>72</sup> This legal standard interprets the term “development” as “referring not merely to augmenting the content generally, but to *materially contributing* to its alleged unlawfulness.”<sup>73</sup> In one early case, *Roommates*, the Ninth Circuit

---

69. *Id.* (“[W]e trained a range of classifiers on new risk vectors and have incorporated these into our monitoring workflow, enabling us to better enforce our API usage policies.”).

70. *Id.* at 21.

71. 47 U.S.C. § 230(f)(3).

72. *Fair Hous. Council v. Roommates.com, LLC*, 521 F.3d 1157, 1168 (9th Cir. 2008); *Kimzey v. Yelp! Inc.*, 836 F.3d 1263, 1269 n.4 (9th Cir. 2016) (“Our sister circuits have generally adopted *Roommates.Com*’s ‘material contribution’ to activity test.”).

73. *Roommates*, 521 F.3d at 1167–68 (emphasis added).

held that a website became a “developer” of content when it required that users provide specific information to use its service: “[b]y requiring subscribers to provide the information as a condition of accessing its service, and by providing a limited set of pre-populated answers” a website provider becomes “much more than a passive transmitter of information provided by others; it becomes the *developer*, at least in part, of that information.”<sup>74</sup> The *Roommates* court distinguished “material contribution” from the protected use of “neutral tools” to filter and display user-generated content using user-generated criteria, which the court opined would not alone constitute “development” of content.<sup>75</sup> In applying the “material contribution” test, courts recognize a “crucial distinction between, on the one hand, taking actions (traditional to publishers) that are necessary to the display of unwelcome and actionable content and, on the other hand, responsibility for what makes the displayed content illegal or actionable.”<sup>76</sup>

In another early case, *Accusearch*, the Tenth Circuit found that a website’s solicitation and online publication of phone record information constituted “development” of that information, even where the service did not modify the content of the records: “[b]y paying its researchers to acquire telephone records, knowing that the confidentiality of the records was protected by law, it *contributed mightily* to the unlawful conduct of its researchers.”<sup>77</sup> The *Accusearch* court found Accusearch’s contribution to the development of content was “more pronounced than that of *Roommates*,” which had “encouraged users to post offending content.”<sup>78</sup> *Roommates*, *Accusearch*, and their progeny highlight the critical inquiry regarding not only the “creation” of information but also the “development” of pre-existing content created by a third party.<sup>79</sup> They also highlight the importance of focusing on the particular aspect of content that is allegedly harmful, and the materiality of a provider’s contribution to that particular component of the information.

Courts have also addressed Section 230 in the context of whether the use of algorithmic tools by search engines, social networks, and other Internet-based services to promote, recommend, or filter content constitutes “development” of content that could bring a provider outside of Section 230

---

74. *Id.* at 1166 (emphasis added).

75. *Id.* at 1167–69.

76. *Kimzey*, 836 F.3d at 1269 n.4 (citations omitted).

77. Fed. Trade Comm’n v. *Accusearch, Inc.*, 570 F.3d 1187, 1200 (10th Cir. 2009) (emphasis added); *see also id.* (“Indeed, *Accusearch*’s responsibility is more pronounced than that of *Roommates.com*. *Roommates.com* may have encouraged users to post offending content; but the offensive postings were *Accusearch*’s *raison d’etre* and it affirmatively solicited them.”).

78. *Id.*

79. *See* 47 U.S.C. § 230(f)(3) (emphasis added) (defining “information content provider” to include “any person or entity that is responsible, in whole or in part, for the creation or development of information”). *See generally Accusearch*, 570 F.3d 1187.

protection. In a recent federal district court case, *Divino*, the claimant alleged that Google “improperly censored or otherwise interfered” with user videos by using “artificial intelligence algorithms” to “review and regulate video content.”<sup>80</sup> The court found that the complaint failed to sufficiently allege that “those algorithms are themselves content or materially contribute to content.”<sup>81</sup> Likewise, allegations that Google “embed[ed] metadata and other discriminatory information into the video content” posted by others also did not support a plausible inference that Google “create[d] or develop[d] any content.”<sup>82</sup>

In another recent case, *Force*, the Second Circuit considered a claim that Facebook unlawfully assisted Hamas in terrorist attacks coordinated on Facebook by promoting “personalized” content.<sup>83</sup> The court held that Facebook’s use of algorithms to “match” content to users “based on objective factors applicable to any content” did not constitute development of the content where the content was “materially unaltered.”<sup>84</sup> The *Force* court concluded that “[m]erely arranging and displaying others’ content to users of Facebook through such algorithms — even if the content is not actively sought by those users — is not enough to hold Facebook responsible as the ‘develop[er]’ or ‘creat[or]’ of that content.”<sup>85</sup> The *Force* court also considered Facebook’s use of “artificial intelligence to block or remove text that might be advocating for terrorism,” which did not constitute “development.”<sup>86</sup>

In another case, *Dyroff*, the Ninth Circuit found that a website did not materially contribute to the alleged unlawfulness of content by employing tools that promote content through recommendation and notification features: “[t]he recommendation and notification functions helped facilitate this user-to-user communication, but it did not materially contribute . . . to the alleged unlawfulness of the content.”<sup>87</sup> Similarly, the Sixth Circuit in *Dirty World* rejected a call to establish an “encouragement test,” concluding that a

---

80. *Divino Grp. LLC v. Google LLC*, No. 19-04749, 2022 U.S. Dist. LEXIS 180048, at \*18, \*51 (N.D. Cal. Sept. 30, 2022).

81. *Id.*

82. *Id.*

83. *Force v. Facebook, Inc.*, 934 F.3d 53, 70 (2d Cir. 2019).

84. *Id.*

85. *Id.* (citing *Marshall’s Locksmith Serv. v. Google, LLC*, 925 F.3d 1263, 1269–70 (D.C. Cir. 2019)) (“The algorithms take the information provided by Facebook users and ‘match’ it to other users — again, materially unaltered—based on objective factors applicable to any content, whether it concerns soccer, Picasso, or plumbers.”).

86. *Force*, 934 F.3d at 60 (internal citations and quotations omitted) (“Facebook is also experimenting with artificial intelligence to block or remove text that might be advocating for terrorism. When Facebook detects terrorist-related content, it also uses artificial intelligence to identify similar, socially interconnected accounts, content, and pages that may themselves support terrorism.”).

87. *Dyroff v. Ultimate Software Grp., Inc.*, 934 F.3d 1093, 1098–99 (9th Cir. 2019).

provider’s “encouragement” of content did not necessarily rise to the level of a material contribution to the development of content.<sup>88</sup> Likewise, in *Marshall’s Locksmith*, the D.C. Circuit opined that Google’s decision to translate third-party address data into a map pinpoint format did not constitute the “creation” or “development” of information under Section 230.<sup>89</sup> The *Marshall’s Locksmith* court noted that “the underlying information is entirely provided by the third party, and the choice of presentation does not itself convert the search engine into an information content provider.”<sup>90</sup> That court also noted that Google used a “neutral algorithm to make that translation.”<sup>91</sup> This line of cases suggests that the use of tools to review, regulate, filter, and translate third-party content without material alteration does not rise to the level of “development” of content.

One recent Section 230 case, however, signals a potential shift in the “information content provider” test that could expose more provider conduct to liability.<sup>92</sup> In *Vargas*, the claimant alleged that Facebook’s “targeting methods provide tools to exclude women of color, single parents, persons with disabilities and other protected attributes,” such that the claimant was “prevented from having the same opportunity to view ads for housing” as other Facebook users.<sup>93</sup> The court analyzed Facebook’s advertising platform, noting that it allowed advertisers to target specific audiences by including or excluding categories of persons.<sup>94</sup> The court considered that “Facebook created the categories, used its own methodologies to assign users to the categories, and provided simple drop-down menus and toggle buttons to allow housing advertisers to exclude protected categories of persons,” and concluded that “Facebook’s *own actions* ‘contribute[d] materially to the alleged illegality of the conduct.’”<sup>95</sup> The court found Facebook was “a co-developer of content.”<sup>96</sup> The *Vargas* opinion suggests that the use of algorithms to scrape online information could push a claim outside of Section 230 protection: “Facebook’s algorithms nevertheless ascertained that

---

88. *Jones v. Dirty World Entm’t Recordings LLC*, 755 F.3d 398, 414 (6th Cir. 2014) (“Under an encouragement test of development, these websites would lose the immunity under the CDA and be subject to hecklers’ suits aimed at the publisher.”).

89. *Marshall’s Locksmith*, 925 F.3d at 1268–70.

90. *Id.* at 1269 (“Indeed, were the display of this kind of information not immunized, nothing would be: every representation by a search engine of another party’s information requires the translation of a digital transmission into textual or pictorial form.”).

91. *Id.* at 1270.

92. *Vargas v. Facebook, Inc.*, No. 21-16499, 2023 U.S. App. LEXIS 27288, at \*5 (9th Cir. Oct. 13, 2023).

93. *Id.* at \*2.

94. *Id.* at \*5–6.

95. *Id.* at \*7 (emphasis added) (citing *Fair Hous. Council v. Roommates.com, LLC*, 521 F.3d 1157, 1168 (9th Cir. 2008)).

96. *Id.* at \*5.

information from the user’s online activities and allowed advertisers to target ads depending on the characteristic.”<sup>97</sup> The *Vargas* opinion also draws into the question the “neutral tools” test: “[a] patently discriminatory tool offered specifically and knowingly to housing advertisers does not become ‘neutral’ within the meaning of this doctrine simply because the tool is also offered to others.”<sup>98</sup>

Further, growing tension between Section 230 and First Amendment protections could also present “land mines” that disrupt Section 230’s “material contribution” test.<sup>99</sup> For example, a social media platform’s assertion that the algorithmic curation of third-party content constitutes expressive conduct protected by the First Amendment could run contrary to the platform’s assertion that it is *not* materially contributing to the development of that content for purposes of Section 230 protection.<sup>100</sup>

These amorphous and evolving legal standards governing Section 230’s “information content provider” test will soon be applied to generative AI, introducing legal risk for providers that currently rely on Section 230’s bar to liability. Under the current standards, courts could justifiably hold that generative AI systems fall outside of Section 230 protection by “contribut[ing] mightily” to the “development” of a pre-existing third-party dataset, such as an underlying Internet-based dataset.<sup>101</sup> This risk heightens the importance of decisions made in the design phase of a generative AI system, during which data is gathered, prepared, labeled, and potentially modified by a system developer. Thus, it is likely that at least some of a generative AI system’s “own actions” could render it a “co-developer of content.”<sup>102</sup> Conversely, some limited functions of generative AI systems are likely to fall within the protected activity of “arranging and displaying others’ content”<sup>103</sup> or the

---

97. *Id.* at \*8.

98. *Id.* at \*10.

99. *See, e.g.*, Transcript of Oral Argument at 149, *Moody v. NetChoice, LLC*, No. 22-277 (2024) (Barrett, J.) (“[T]here are a bunch of land mines. And if that’s a land mine, if what we say about this is that this is speech that’s entitled to First Amendment protection, I do think then that has Section 230 implications for another case, and so it’s always tricky to write an opinion when you know there might be land mines that would affect things later.”).

100. *Id.* at 122 (Gorsuch, J.) (“Section 230 says we’re not going to treat you as publishers so long as you are not — it’s not your communication in whole or in part is what the definition says. And if it’s now their communication in part, do they lose their 230 protections?”); *id.* (Gorsuch, J.) (“[W]hy isn’t it their communication in part if it — if it’s part of this larger mosaic of editorialized discretion and the whole feel of the website?”).

101. *Fed. Trade Comm’n v. Accusearch, Inc.*, 570 F.3d 1187, 1200 (10th Cir. 2009).

102. *Vargas*, No. 21-16499, 2023 U.S. App. LEXIS 27288, at \*5–7.

103. *Force v. Facebook, Inc.*, 934 F.3d 53, 70 (2d Cir. 2019) (citing *Marshall’s Locksmith Serv. v. Google, LLC*, 925 F.3d 1263, 1269–70 (D.C. Cir. 2019)) (“The algorithms take the information provided by Facebook users and ‘match’ it to other users—again, materially unaltered—based on objective factors applicable to any content, whether it concerns soccer, Picasso, or plumbers.”).

protected “translation of a digital transmission into textual or pictorial form.”<sup>104</sup>

Ultimately, whether Section 230 applies to generative AI systems may turn on several technical and fact-intensive considerations regarding the manner and extent of an AI system’s contribution to the creation or development of AI output content. The design, development, and deployment of an AI system should be considered in turn. Relevant attributes of an AI system’s design include data strategy, selection, and preparation; the relevant attributes of an AI system’s development include AI model selection, training, tuning, and testing; the relevant attributes of an AI system’s post-deployment measures include model retraining, tuning, and maintenance. Other considerations include the source (e.g., model developer, deployer, user) of model training, tuning, RLHF, and human intervention, and the extent to which these contribute to the harmfulness of AI output content; the extent to which AI output content is an extraction of third-party content from training data; the extent to which an AI system acts as a mere intermediary between a user and a pre-existing dataset, i.e., whether the application merely retrieves information in response to a prompt; the extent to which AI output content is a “hallucination”; the role of the user, including prompting and instruction; whether a user “jailbreaks” an AI system to generate harmful content that the system has been trained not to generate; and the specific complained-of aspect of the content and the extent to which each of the various components of an AI system materially contributed to its creation or development. Regardless of the legal standards applied, generative AI systems are unlikely to categorially evade classification as an “information content provider” under current Section 230 jurisprudence, introducing substantial legal risk in the development of generative AI systems.

*B. Whether Generative AI Systems Fall Outside the Scope of Section 230  
“Publisher or Speaker” Protection*

Section 230 protects a provider of an interactive computer service from claims that attempt to hold the provider liable as a “publisher or speaker” of third-party content.<sup>105</sup> This immunity generally extends to claims based on the “exercise of a publisher’s traditional editorial functions — such as deciding whether to publish, withdraw, postpone, or alter content.”<sup>106</sup> As one court

---

104. *See Marshall’s Locksmith*, 925 F.3d at 1268–70 (“every representation by a search engine of another party’s information requires the translation of a digital transmission”).

105. 47 U.S.C. § 230(c)(1).

106. *McCall v. Zotos*, No. 22-11725, 2023 U.S. App. LEXIS 14585, at \*8 (11th Cir. June 12, 2023) (citing *Zeran v. Am. Online, Inc.*, 129 F.3d 327, 330 (4th Cir. 1997)); *see also HomeAway.com v. City of Santa Monica*, 918 F.3d 676, 681 (9th Cir. 2018) (“Although the CDA does not define ‘publisher,’

opined, “the question is whether a plaintiff’s claim arises from a third party’s information, and — crucially — whether to establish the claim the court must necessarily view the defendant, not as a publisher in the abstract, but rather as *the* publisher of that third-party information.”<sup>107</sup> Section 230 applies not only to defamation claims, where publication is an explicit element of the offenses, but also to claims where “the duty that the plaintiff alleges the defendant violated derives from the defendant’s *status or conduct* as a publisher or speaker.”<sup>108</sup> For example, a social network may invoke Section 230 immunity in a claim that it failed to remove a harmful or threatening profile because a decision regarding whether to retract content is generally protected publishing activity.<sup>109</sup> Courts have invoked “the prophylaxis of section 230(c)(1) in connection with a wide variety of causes of action, including housing discrimination, negligence, and securities fraud and cyberstalking.”<sup>110</sup> But courts disagree on the exact scope of protected publisher activity and diverge on whether Section 230 immunity extends to preclude claims for negligent design, product liability, and distributor liability.

Some cases have construed protected “publisher or speaker” activity broadly. In *Backpage*, the First Circuit opined that a website’s decisions regarding what terms were allowed in advertising postings were publisher functions because decisions like these “reflect choices about what content can appear on the website and in what form,” and are thus “editorial choices that fall within the purview of traditional publisher functions.”<sup>111</sup> Similarly, the Second Circuit in *Force* held that Facebook’s use of advanced algorithms to promote and filter “personalized” content constituted an editorial decision regarding how and where to display content, which was within the scope of publisher activity under Section 230.<sup>112</sup> The Fifth Circuit in *Doe v. Myspace* applied Section 230 to immunize MySpace from a claim that it failed to implement basic safety measures to prevent sexual predators from communicating with minors on its platform.<sup>113</sup> The *Myspace* court found that Section 230 barred the claims against Myspace “notwithstanding plaintiff’s assertion that they only [sought] to hold MySpace liable for its failure to implement measures that would have prevented [the communication],” which

---

we have defined ‘publication’ in this context to ‘involve[] reviewing, editing, and deciding whether to publish or to withdraw from publication third-party content.’” (citation omitted).

107. *Force*, 934 F.3d at 81 (2d Cir. 2019) (emphasis original).

108. *Barnes v. Yahoo!, Inc.*, 570 F.3d 1096, 1102 (9th Cir. 2009) (emphasis added).

109. *Fed. Trade Comm’n v. LeadClick Media, LLC*, 838 F.3d 158, 174 (2d Cir. 2016) (citations and internal quotations omitted).

110. *Jane Doe No. 1 v. Backpage.com, LLC*, 817 F.3d 12, 19 (1st Cir. 2016) (internal citations omitted).

111. *Backpage*, 817 F.3d at 20–21.

112. *Force*, 934 F.3d at 66–67.

113. *Doe v. MySpace, Inc.*, 528 F.3d 413 (5th Cir. 2008).



the court described as a veiled attempt to hold Myspace “liable for publishing the communications.”<sup>114</sup> The Sixth Circuit in *Fastcase* applied Section 230 publisher immunity to “automated editorial acts” such as “removing spaces and altering font.”<sup>115</sup> The Ninth Circuit in *Carafano* applied Section 230 to shield a website’s “decision to structure the information provided by users” to allow for “matching” user profiles with similar characteristics.<sup>116</sup> The Ninth Circuit broadly held that any claim that seeks to require the editing or removal of third-party content would constitute an impermissible attempt to hold the service liable as a “publisher or speaker.”<sup>117</sup>

Other cases suggest a trend in Section 230 jurisprudence that may narrow the legal standards governing the “publisher or speaker” test. In early 2024, a California federal district court in *Ziencik*, an online bullying case, found that a claim that Snap breached a duty to provide information to law enforcement had “nothing to do with Snap’s editing, monitoring, or removing of the content” and was “focused squarely on Snap’s *own conduct*” such that the claim did not attempt to hold Snap liable as a “publisher or speaker.”<sup>118</sup> In *Lemmon*, the Ninth Circuit, which had previously interpreted Section 230 to permit some design defect claims to proceed, allowed a claim against Snap based on the “predictable consequences of *designing* Snapchat in such a way that it allegedly encourage[d] dangerous behavior.”<sup>119</sup> The *Lemmon* court endorsed the concept that “Congress has not provided an all-purpose get-out-of-jail-free card for businesses that publish user content on the internet, though any claims might have a marginal chilling effect on internet publishing businesses.”<sup>120</sup> The Tenth Circuit in *Accusearch* noted that Section 230 immunity does not apply where a claim seeks to hold a provider “liable for its *conduct* rather than for the *content* of the information it was offering.”<sup>121</sup> Similarly, the Second Circuit in *LeadClick* permitted a claim that held a provider “accountable for its *own* deceptive acts or practices — for directly

---

114. *Id.* at 420.

115. *O’Kroley v. Fastcase, Inc.*, 831 F.3d 352, 355 (6th Cir. 2016).

116. *Carafano v. Metrosplash.com, Inc.*, 339 F.3d 1119, 1124–25 (9th Cir. 2003).

117. *Google LLC v. Equustek Sols. Inc.*, No. 17-04207, 2017 U.S. Dist. LEXIS 182194, at \*6–7 (N.D. Cal. Nov. 2, 2017) (“The Ninth Circuit has held that, regardless of the underlying cause of action, a claim treats an intermediary as a publisher when it requires the intermediary to remove third-party content”); *Bride v. Snap Inc.*, No. 21-06680, 2023 U.S. Dist. LEXIS 5481, at \*19 (C.D. Cal. Jan. 10, 2023) (“[T]he court finds that Plaintiffs’ theory would require the editing of third-party content, thus treating Defendants as a publisher of content.”).

118. *Ziencik v. Snap, Inc.*, No. 21-7292, 2024 U.S. Dist. LEXIS 12105, at \*4–6 (C.D. Cal. Jan. 19, 2024) (emphasis added).

119. *Lemmon v. Snap, Inc.*, 995 F.3d 1085, 1094 (9th Cir. 2021) (citation and internal quotation omitted) (emphasis added) (holding that Snap is not entitled to immunity under Section 230).

120. *Id.* at 1094 (citing *Internet Brands*, 824 F.3d at 853).

121. *Fed. Trade Comm’n v. Accusearch, Inc.*, 570 F.3d 1187, 1204–05 (10th Cir. 2009) (Tymkovitch, J., concurring) (emphasis in original).

participating in the deceptive scheme by providing edits to affiliate webpages,” which the court found was “not derived from [the provider’s] status as a publisher or speaker.”<sup>122</sup>

A recent dissenting opinion from the Fifth Circuit argued that the “publisher or speaker” standard has become overbroad, positing that Section 230 does not preclude claims based on design defect, product liability, or distributor liability.<sup>123</sup> The Fifth Circuit had initially found that Section 230 provided immunity to Snap in a claim based on a teacher’s use of Snapchat to send sexually explicit material to a minor.<sup>124</sup> The Fifth Circuit later denied rehearing *en banc*, with seven judges dissenting from the denial of rehearing on the basis that the court’s opinion on the scope of Section 230 protection provided “sweeping immunity for social media companies that the text [of Section 230] cannot possibly bear.”<sup>125</sup> The dissent instead argued that “[i]mmunity from design defect claims is neither textually supported nor logical because such claims fundamentally revolve around the platforms’ conduct, not third-party conduct,”<sup>126</sup> and that “[p]roduct liability claims do not treat platforms as speakers or publishers of content.”<sup>127</sup> The dissent further pushed to overturn the Fifth Circuit’s existing “broad-based immunity, including against design defect liability and distributor liability.”<sup>128</sup>

The unresolved and evolving legal standards under Section 230’s “publisher or speaker” test present significant risk regarding whether the novel design and functionality of generative AI systems could give rise to legal claims based on design, distributor, and product liability outside of protected publisher activity. As one court commented, “[t]he broad construction accorded to section 230 as a whole has resulted in a *capacious* conception of what it means to treat a website operator as the publisher or speaker of information provided by a third party.”<sup>129</sup> While some limited generative AI system functions could arguably fall within “traditional editorial functions,”<sup>130</sup> such as the protected use of algorithms to filter and

---

122. Fed. Trade Comm’n v. LeadClick Media, LLC, 838 F.3d 158, 176 (2d Cir. 2016); *see also* Transcript of Oral Argument at 148, *Moody v. NetChoice, LLC*, No. 22-277 (2024) (Solic. Gen. Prelogar) (“[I]f the thing that’s causing harm is the platform’s own conduct in how it structures its service, that’s something that might not be immunized under Section 230.”).

123. *Doe v. Snap, Inc.*, No. 22-20543, 2023 U.S. App. LEXIS 33501 (5th Cir. Dec. 18, 2023) (Elrod, J., dissenting).

124. *Doe v. Snap, Inc.*, No. 22-20543, 2023 U.S. App. LEXIS 16095 (5th Cir. June 26, 2023).

125. *Doe v. Snap, Inc.*, 2023 U.S. App. LEXIS 33501 at \*2.

126. *Id.* at \*5.

127. *Id.*

128. *Id.* at \*4 (citing *Doe v. MySpace, Inc.*, 528 F.3d 413, 421 (5th Cir. 2008)).

129. *Jane Doe No. 1 v. Backpage.com, LLC*, 817 F.3d 12, 19 (1st Cir. 2016) (emphasis added).

130. *HomeAway.com v. City of Santa Monica*, 918 F.3d 676, 681 (9th Cir. 2018).

display “personalized” content based on user inputs,<sup>131</sup> generative AI systems could be exposed to negligent design or product liability claims grounded in a system’s own “conduct” rather than the information contained in third-party training data.<sup>132</sup> Even if a claim is purportedly based on content substantially derived from a third-party dataset used to train an AI system, the system provider’s role in system design and development — including data selection, preparation, and labelling, and model selection, engineering, and training — may constitute “conduct” that falls outside of protected Section 230 publisher activity.<sup>133</sup>

The evolving standards of Section 230 indicate that generative AI systems are unlikely to enjoy unfettered Section 230 immunity as protected “publishers” or “speakers” of third-party content in claims based on the novel design and functionality of generative AI systems. Courts have yet to clarify the scope of the “publisher or speaker” test as applied to generative AI systems, and their opinions will impact Section 230’s broader application not only to AI but to all Internet-based interactive computer services.

## V. CONCLUSION

Section 230 of the Communications Decency Act, the foundational law that catalyzed the development of the Internet by providing legal protections to Internet-based computer services, has reached a breaking point. U.S. courts have allowed the legal standards governing Section 230’s expansive protections to evolve alongside technology, but in doing so, have also stretched Section 230 to its logical bounds. Section 230’s text establishes a bright line that courts will not cross: Section 230 immunity will not apply to technologies that materially contribute to the creation or development of harmful content. Generative AI will thus present courts with novel cases that require application of decades-old Section 230 legal standards to increasingly ubiquitous and self-improving technology that generates content. Courts will be presented with cases that will require them to examine how AI system design and development, user prompts, training data, models, algorithms, and machine learning contribute to generative AI outputs, and the extent to which these components render a generative AI system an “information content provider” or a “publisher or speaker” under Section 230. Judicial opinions on generative AI and Section 230 will have a sweeping impact on the future

---

131. *See Force v. Facebook, Inc.*, 934 F.3d 53, 66–67 (2d Cir. 2019).

132. *See Fed. Trade Comm’n v. LeadClick Media, LLC*, 838 F.3d 158, 176 (2d Cir. 2016); *Ziencik v. Snap, Inc.*, No. 21-7292, 2024 U.S. Dist. LEXIS 12105, at \*4–6 (C.D. Cal. Jan. 19, 2024).

133. *Id.*; *see Lemmon v. Snap, Inc.*, 995 F.3d 1085, 1094 (9th Cir. 2021) (permitting a claim based on the “predictable consequences of designing” a platform).

development of Internet-based technologies. At stake is whether the legal protections that catalyzed the Internet's development prior to the introduction of generative AI will do the same to maintain a "vibrant and competitive free market" for generative AI and the future of the Internet.<sup>134</sup> Based on current jurisprudence delineating the outer bounds of Section 230 immunity, the foundational legal protections underpinning the Internet are unlikely to extend with the same force to generative AI systems and the new frontier of Internet technology.

---

134. 47 U.S.C. § 230(b)(2).