# Generative AI Meets Section 230: The Future of Liability and Its Implications for Startup Innovation

*Megan E. Cistulli*[*]

*The rapid advancement of generative artificial intelligence (AI) is testing the limits of Section 230 of the Communications Decency Act, a statute that has long shielded online platforms from liability for user-generated content. This liability shield helped shape the modern internet, but AI's ability to create its own content blurs the traditional distinction between platforms acting as passive hosts and those functioning as active publishers. As a result, courts and lawmakers are reexamining the scope and future of Section 230. This Comment examines how proposed reforms to Section 230 could impact startups and emerging tech companies that use generative AI in their products and services. It argues that broad rollbacks or carve-outs from Section 230 protections would impose disproportionate burdens on smaller companies, exposing them to increased litigation risks, major compliance costs, and crucially, barriers to innovation. Using a comparative analysis of existing proposals, caselaw, and international AI regulations, this Comment introduces a proportional framework that combines risk-based tiers with sandbox innovation and scales obligations to company capacity while fostering responsible experimentation. The framework's goal is to hold companies accountable without stifling innovation for smaller market entrants. As a result, this model offers a balanced path forward for regulating generative AI content in the digital age.*



[*] J.D., 2026, University of Chicago Law School; M.B.A., 2026, University of Chicago Booth School of Business; B.A., 2022, University of California, Berkeley. I am grateful to the staff of *The University of Chicago Business Law Review* for their insightful feedback through many rounds of drafting. Equally important, I thank Professor Vincent Buccola for working with me to shape and structure this piece in a way that bridges the practical realities of business with the theoretical foundations of the law. Last but not least, thank you to my family and to my partner, Andriana Acosta, for their immeasurable support.

## I. INTRODUCTION

Courts have a novel question in front of them: does Section 230 of the Communications Decency Act protect AI-generated content?[1]

Section 230 protects online platforms from legal responsibility for content created by others.[2] This foundational law ensures websites, social media companies, and other online services are not treated as "publishers" or "speakers" of user-generated content. The Act's protection continues to play an instrumental role in fostering innovation, enabling free expression, and shaping the modern internet.[3] At the same time, it sparks ongoing debates about accountability, content moderation, and the evolving role of

---

1 The Telecommunications Act of 1996 included the Communications Decency Act, Pub. L. No. 104-104, § 509, 110 Stat. 56 (1996), codified as amended at 47 U.S.C. § 230 (2018) (hereinafter Section 230).

2 *Id*.

3 *Id*.

digital platforms.[4] Courts have traditionally stretched Section 230[5] to adapt to new technologies, but there is a limit: platforms lose immunity if they significantly help create harmful content.[6] Key legal tests, such as the material contribution test[7] and the neutral tools test,[8] have helped courts determine when Section 230 immunity applies and when it does not. Generative artificial intelligence (AI) creates unique challenges because it generates its own content based on prompts, training data, and algorithms. Courts will have to decide whether these AI systems count as "content creators" or just neutral platforms. These decisions will redefine the legal landscape for startup technology companies leveraging generative AI, and they will influence content moderation practices, liability exposure, and operational costs while also defining the balance between innovation and accountability needed to sustain the generative AI ecosystem.

Generative AI is not just transforming industries; it is reshaping the very concept of user creation.[9] Unlike predictive AI, which is designed to forecast, classify, or score outcomes, such as predicting creditworthiness, detecting fraud, or recommending a next purchase, generative AI produces new, original outputs: text, images, videos, music, and more, which blurs the lines between human ingenuity and algorithmic outputs. Yet, as this revolutionary technology dazzles with its capabilities, it also poses perplexing legal questions: when machines create, whose voice is it?[10] Is it the user's speech, the platform's creation, or something entirely new? Can the law[11] that once "created the Internet"[12] shield these AI-generated outputs, or has it reached its limit?[13] More

---

4 *Id.*; *see also* ERIC N. HOLMES, CONG. RSCH. SERV., R47753, LIABILITY FOR ALGORITHMIC RECOMMENDATIONS (2023).

5 *See* Zeran v. Am. Online, Inc., 129 F.3d 327 (4th Cir. 1997).

6 Peter J. Benson & Valerie C. Brannon, CONG. RESEARCH SERV., IF12584, SECTION 230: A BRIEF OVERVIEW (2024).

7 *See* Fair Hous. Council of San Fernando Valley v. Roommates.com, LLC, 521 F.3d 1157 (9th Cir. 2008).

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

9 Shalwa, *Generative AI in 2024: Market Growth, Industry Impact, and Key Statistics*, ARTSMART.AI (Aug. 9, 2024), https://perma.cc/8TNB-2TXU; *see also* QuantumBlack, AI by McKinsey, *The State of AI in Early 2024: Gen AI Adoption Spikes and Starts to Generate Value*, MCKINSEY & COMPANY (2024), https://perma.cc/2E7E-HP5U; *The Elastic Generative AI Report*, ELASTIC (Mar. 2024), https://perma.cc/96J9-C9AS.

10 Eugene Volokh, Mark A. Lemley & Peter Henderson, *Freedom of Speech and AI Output*, 3 J. FREE SPEECH L. 651 (2023).

11 47 U.S.C. § 230 (c)(1) (2018).

12 *See generally* Jeff Kosseff, The Twenty-Six Words That Created the Internet (2019).

13 Graham H. Ryan, Generative AI Will Break the Internet: Beyond Section 230, 37 HARV. J.L. & TECH. (Spring 2024).

importantly, and the central question answered in this article: how would the removal or narrowing of Section 230[14] immunity affect startups and early-stage businesses that integrate generative AI into their products or services?

Applied expansively by U.S. courts, Section 230's[15] purpose is "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."[16] Enacted in 1996, Section 230[17] catalyzed the internet's growth. Its text, however, remains stagnant while the internet experienced a user growth rate of over 13,000%; the internet boomed from 40 million users[18] to over 5 billion.[19] At the time of Section 230's[20] inception, the law granted protection to passive platforms like AOL's online message boards.[21] In the modern day, courts continue to construe Section 230[22] broadly, where courts afford protection to platforms like Facebook when harmful content is shared on their platforms by users.[23] Similarly, major technology companies, such as Google, are not responsible for simply grabbing and subsequently displaying harmful content already online so long as Google did not originate the material.[24]

The rise of AI coincides with increasing scrutiny of Section 230.[25] Recent legislative proposals, including a 2024 initiative to sunset Section 230 by 2025,[26] raise significant concerns about the implications for startups relying on generative AI since Section 230 shields "interactive computer services" from liability for third-party content.[27] As startups integrate AI-driven content

---

14 47 U.S.C. § 230.

15 *Id*. § 230(c)(1).

16 *Id*. § 230(b)(2).

17 *Id*. § 230(c)(1).

18 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)).

19 *See* Ani Petrosyan, *Number of Internet and Social Media Users Worldwide as of October 2024*, STATISTA (Nov. 4, 2024), https://perma.cc/TKR3-58H5 ("As of October 2024, there were 5.52 billion internet users worldwide, which amounted to 67.5 percent of the global population. Of this total, 5.22 billion, or 63.8 percent of the world's population, were social media users.").

20 *Id*. § 230.

21 *Zeran*, 129 F.3d at 328.

22 47 U.S.C. § 230.

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

24 *See* Marshall's Locksmith Serv. v. Google, LLC, 925 F.3d 1263 (D.C. Cir. 2019).

25 HOLMES, *supra* note 4.

26 Legislative Proposal to Sunset Section 230 of the Communications Decency Act: Hearing Before the H. Comm. on Energy & Com., 118th Cong. (May 22, 2024).

27 47 U.S.C. § 230.

generation into their operations, changes to Section 230 or exclusions from its protections could amplify legal risks, impose greater compliance demands, and heighten operational expenses, which poses unique challenges for smaller tech firms.

This Comment explores how proposed legislative changes to Section 230 will reshape the legal landscape for startups integrating generative AI into their business models, which is a significant portion of them.[28] It argues that removing or significantly narrowing Section 230 protections for AI-generated content would impose heightened liability risks, compliance challenges, and operational costs on emerging companies. To balance accountability with innovation, this Comment advocates for a framework that preserves competition while ensuring responsible AI deployment. Using a comparative analysis of state-specific regulations and key caselaw, it examines how changes to Section 230 could impact content moderation, liability exposure, and business sustainability. Ultimately, this Comment advances a proportional regulatory framework that tailors obligations to both company capacity and the risk level of the AI system while embedding sandbox innovation as a tool for safe experimentation with the central aim of fostering an ecosystem that encourages innovation while ensuring meaningful accountability.

This Comment proceeds as follows: Part II examines the foundational role of Section 230 in shaping internet governance and its evolving relevance to AI-generated content. Part III builds on this foundation by analyzing key legal precedents and proposed legislative reforms that could redefine liability for AI-driven platforms while assessing how these changes might alter platform responsibility and legal exposure. Shifting to the practical implications of these reforms, Part IV explores business models that rely on Section 230 protections by highlighting the potential disruptions that regulatory changes may introduce. Part V evaluates the financial and operational burdens these changes could impose on startups; at the same time, it emphasizes the disproportionate challenges faced by emerging AI companies that may struggle to absorb heightened compliance costs. To address these concerns, Part VI proposes a proportional regulatory framework that pairs risk-based tiers with sandbox innovation and ensuring that obligations scale to company capacity while creating a safe space for startups to experiment without undermining

---

[28] *See generally* FENWICK & WEST LLP, VENTURE BEACON Q1 2025 REPORT (2025), https://perma.cc/VUD8-KM7Z.

accountability. Finally, Part VII concludes by advocating for reforms that promote responsible AI development while preserving competition and enabling technological progress.

## II. GENERAL BACKGROUND ON AI & SECTION 230

### A. Generative AI Explained

Generative artificial intelligence is a revolutionary branch of AI technology that redefined content creation by enabling systems to produce entirely new material, including text, images, audio, and video. Unlike traditional AI systems, which primarily analyze or classify existing data, generative AI models use advanced algorithms, often trained on vast datasets, to generate novel outputs that mimic human creativity and problem-solving abilities.[29] This paradigm shift presents profound implications for industries ranging from entertainment and journalism to education and legal services.

At the core of generative AI are large language models (LLMs) and other foundational models. These systems, such as OpenAI's GPT series, function by predicting the most likely continuation of input data based on patterns identified during training. For example, GPT-4, a leading LLM, processes extensive datasets encompassing text from books, articles, and websites to craft coherent and contextually relevant responses to user prompts.[30] Other notable examples include DALL-E,[31] which generates images from textual descriptions, and generative models used for audio synthesis and video creation.

The development of generative AI systems involves multiple stages including the collection and preprocessing of training data, the selection and fine-tuning of machine learning algorithms, and reinforcement learning from human feedback (RLHF). These models are designed to improve over time and incorporate user interactions and additional datasets to refine their outputs. However, this adaptability also introduces unique challenges, such as the risk of perpetuating biases present in the training data or generating inaccurate or harmful content.

The rise of generative AI has sparked considerable debate regarding its regulatory and ethical implications across disciplines

---

29 Ian Goodfellow et al., DEEP LEARNING (MIT Press 2016).

30 OpenAI et al., *GPT-4 Technical Report*, CORNELL UNIV., https://perma.cc/5QGW-MC4P (Mar. 15, 2023, last revised Mar. 4, 2024).

31 OpenAI, *DALL-E: Creating Images from Text*, OpenAI Blog (Jan. 5, 2021), https://perma.cc/L77W-7EAN.

from copyright to defamation.[32] These systems challenge existing legal frameworks, particularly Section 230,[33] which has historically shielded online platforms from liability for third-party content. Generative AI's capacity to independently create content raises questions about whether these systems qualify as "information content providers" under Section 230, potentially excluding them from its protections.[34]

1. Section 230's Legal History & Significance

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."[35] It is foundational in enabling platforms to host user-generated content. Section 230[36] historically played a pivotal role in shaping the internet by enabling platforms to host user-generated content with limited liability. This legal provision shields internet service providers and online platforms from accountability for content created by their users, such as YouTube, which allows users to upload and share their own content. Its inception was influenced by the contrasting rulings in *Cubby, Inc. v. CompuServe, Inc.*, where the court declined to hold CompuServe liable as a publisher of defamatory content, and *Stratton Oakmont, Inc. v. Prodigy Servs. Co.*, which found Prodigy liable due to its content moderation practices.[37] These conflicting outcomes prompted the creation of Section 230[38] and provided a legal framework encouraging innovation and growth in the burgeoning tech sector.[39]

The Ninth Circuit acknowledged the critical role of Section 230[40] in limiting litigation against online platforms. In describing the potential consequences of removing those protections, the

---

[32] Micaela Mantegna, *ARTificial: Why Copyright Is Not the Right Policy Tool to Deal with Generative AI*, 133 YALE L.J.F. 1126 (Apr. 22, 2024); *see also* Walters v. OpenAI, LLC, No. 23-A-04860-2 (Ga. Super. Ct. May 19, 2025).

[33] 47 U.S.C. § 230.

[34] *Id*. § 230(c)(1).

[35] *Id*.

[36] *Id*. § 230.

[37] *Compare* Cubby, Inc. v. CompuServe, Inc., 776 F. Supp. 135 (S.D.N.Y. 1991) *with* Stratton Oakmont, Inc. v. Prodigy Servs. Co., No. 31063, 1995 WL 323710 (N.Y. Sup. Ct. May 24, 1995).

[38] 47 U.S.C. § 230.

[39] Eric Goldman, *Why Section 230 Is Better Than the First Amendment*, 95 NOTRE DAME L. REV. REFLECTION 33 (2019).

[40] 47 U.S.C. § 230.

court warned of "death by ten thousand duck-bites"[41]: a vivid metaphor for the flood of lawsuits platforms might face without immunity. For startups in particular, legal exposure at that level could quickly overwhelm their limited resources, which could stifle their innovation and growth.[42] This sentiment highlights the importance of Section 230[43] in allowing smaller companies to operate without the looming threat of excessive legal exposure. Section 230[44] also plays a pivotal role in fostering a competitive and innovative tech marketplace by providing liability protections that encourage product development and consumer choice. Weakening or removing these protections would disproportionately impact smaller businesses and new market entrants who often lack the resources to manage increased litigation and compliance costs.[45] Such changes could consolidate power among larger, established companies thereby reducing competition and limiting opportunities for innovation. Over time, the scope of Section 230's protections expanded through landmark cases. In *Zeran v. Am. Online, Inc.*,[46] the court upheld broad immunity for online platforms from liability for user-generated content, even in circumstances involving harmful material. Similarly, *Carafano v. Metrosplash.com, Inc.*[47] reinforced this expansive interpretation and solidified Section 230's role as a shield for platforms hosting third-party content.

2. Key Cases Shaping Section 230 Immunity

A series of key cases illustrates how reforms to Section 230 could reshape these legal protections and potentially increase liability risks and operational costs for tech companies. In *Barnes v. Yahoo!, Inc.*,[48] the Ninth Circuit clarified the limits of Section

---

41 Fair Hous. Council of San Fernando Valley v. Roommates.com, LLC, 521 F.3d 1157, 1174 (9th Cir. 2008)

42 *See* Evan Engstrom, *Primer: Value of Section 230*, ENGINE (Jan. 31, 2019), https://perma.cc/V2K6-HKB5 (breaking down the cost of litigation for startup and early-stage companies).

43 47 U.S.C. § 230.

44 *Id*.

45 Jennifer Huddleston, *Competition and Content Moderation: How Section 230 Enables Increased Tech Marketplace Entry*, POL'Y ANALYSIS NO. 922 (2022); *see also* Ryan Nabil, *Why Repealing Section 230 Will Hurt Startups and Medium-Sized Online Businesses*, COMPETITIVE ENTERPRISE INST. (Feb. 1, 2021), https://perma.cc/7K74-V2FQ; Cory Doctorow, *Wanna Make Big Tech Monopolies Even Worse? Kill Section 230*, ELECTRONIC FRONTIER FOUND., (May 24, 2024), https://perma.cc/B8H4-DVSG.

46 Zeran v. Am. Online, Inc., 129 F.3d 327 (4th Cir. 1997).

47 Carafano v. Metrosplash.com, Inc., 339 F.3d 1119 (9th Cir. 2003).

48 Barnes v. Yahoo!, Inc., 570 F.3d 1096 (9th Cir. 2009).

230 immunity by distinguishing between publisher-related protections and other liabilities like breach of contract. Yahoo was held liable for failing to honor a promise to remove harmful content. The case illustrates that if Section 230 protections are narrowed, companies could face increased exposure to claims that fall outside traditional publisher liability—a risk that could disproportionately burden startups lacking the resources for extensive legal defenses. Moreover, startups may face pressure to shoulder heavy responsibilities in content oversight, customer service, and data privacy far earlier than their stage of growth can reasonably support.[49]

The case of *Gonzalez v. Google LLC* highlights potential liability for algorithmic promotion of harmful content.[50] The parents of a victim in an ISIS attack sued Google alleging liability for promoting terrorist content through its platform.[51] Similarly, in *Twitter, Inc. v. Taamneh*, the United States Supreme Court addressed allegations that Twitter, Facebook, and Google knowingly allowed ISIS to use their platforms.[52] The Court ruled that liability under the Antiterrorism Act requires clear evidence of aiding and abetting specific acts of terrorism.[53] These cases collectively highlight the challenges of holding platforms accountable for user-generated content with the Court in *Gonzalez* leaving Section 230 protections intact and deferring broader considerations to future litigation.[54] For smaller tech companies, this could mean significant expenses to overhaul recommendation and moderation systems.

In another case, *Lemmon v. Snap, Inc.*, the Ninth Circuit ruled that Snap could be sued over its "speed filter" design, finding Section 230 immunity does not shield negligent product design even when the feature is a so-called neutral tool.[55] For startups, the case highlights how narrowing Section 230 could expose them to costly design-based lawsuits even when they are not directly responsible for user content. In *Federal Trade Commission v. Accusearch Inc.*,[56] the Tenth Circuit denied Section 230 protections to Accusearch for materially contributing to unlawful

---

49 LING ZHU & LAURIE HARRIS, CONG. RSCH. SERV., R47569, GENERATIVE ARTIFICIAL INTELLIGENCE AND DATA PRIVACY: A PRIMER (2023).

50 Gonzalez v. Google LLC, 2 F.4th 871 (9th Cir. 2021).

51 *Id*.

52 Twitter v. Taamneh, 598 U.S. 471 (2023).

53 18 U.S.C. §§ 2331–2339D (2023).

54 *Google LLC*, 2 F.4th 871.

55 Lemmon v. Snap, Inc., 995 F.3d 1085 (9th Cir. 2021).

56 Federal Trade Commi. v. Accusearch Inc., 570 F.3d 1187 (10th Cir. 2009).

conduct by selling private data.[57] This case underscores that reforms could require startups to enhance oversight over the content they generate or facilitate, significantly increasing operational costs especially for generative AI companies. Along a similar vein, *FTC v. LeadClick Media, LLC*[58] demonstrates that Section 230[59] does not protect companies that actively engage in illegal conduct. LeadClick's role in misleading advertising led to liability and signals that reforms could necessitate rigorous content monitoring for platforms enabling third-party content, which adds operational burdens for startups.

*a) Material Contribution Test: When Platforms Shape Content*

Courts have ruled that platforms lose Section 230 immunity when they play an active role in developing or shaping unlawful content. This material contribution test is central in *Fair Housing Council v. Roommates.com*,[60] where the Ninth Circuit held that Roommates.com was liable for requiring users to answer discriminatory housing preference questions. By structuring its platform to facilitate illegal conduct, the company was deemed responsible for the content. However, the court distinguished this from an open-text field where users could freely describe their preferences; since Roommates.com did not dictate those responses, it remained immune under Section 230[61] for that content. Similarly, in *FTC v. Accusearch Inc.*,[62] a website that sold confidential phone records could not claim Section 230 immunity. The court found that the company actively solicited, paid for, and knowingly distributed illegally obtained data, so it makes them an information content provider rather than a neutral platform. These cases demonstrate that platforms cannot hide behind Section 230 when they materially contribute to illegal content or facilitate its creation.

---

57 *Id*.

58 FTC v. LeadClick Media, LLC, 838 F.3d 158 (2d Cir. 2016).

59 47 U.S.C. § 230.

60 Fair Hous. Council of San Fernando Valley v. Roommates.com, LLC, 521 F.3d 1157 (9th Cir. 2008).

61 47 U.S.C. § 230.

62 *Accusearch Inc.*, 570 F.3d 1187.

*b) Neutral Tools Test: Passive vs. Active Content Moderation*

Courts have also recognized that platforms using neutral tools, specifically features designed for general use rather than targeting specific illegal activity, remain protected under Section 230. This principle is reinforced in *O'Kroley v. Fastcase*,[63] where a plaintiff sued Google after its search algorithm displayed misleading snippets that falsely suggested he had committed a crime. The court ruled that Google's search function merely rearranged existing content rather than creating new material which means Section 230 shielded the company from liability.[64]

A similar approach was taken in *Force v. Facebook*, where the Second Circuit found that Facebook's algorithmic content recommendations did not make the company liable for terrorist propaganda on its platform.[65] Because the algorithm applied neutral ranking principles to all content, rather than specifically promoting harmful material, Facebook retained immunity. These rulings confirm that automated tools, even if they influence content visibility, do not necessarily strip platforms of Section 230 protection unless they actively shape or develop the content.

Legal scholars suggest that AI models exist on a spectrum: some function like search engines, retrieving and summarizing existing data, potentially covered by Section 230, while others generate original content based on predictive algorithms more akin to a publisher.[66] In *Accusearch*, the court denied immunity because the company procured and distributed illegal information. If AI tools actively generate false or unlawful material rather than merely retrieving or organizing data, they could face similar liability under future rulings.[67]

*c) Recent Caselaw*

The 2025 case of *Walters v. OpenAI*, though not a Section 230 dispute, offers an early and important glimpse into how courts may allocate liability when generative AI is involved.[68] The case

---

63 O'Kroley v. Fastcase Inc., No. 3:13-0780 (M.D. Tenn. May. 27, 2014) (granting Google's motion to dismiss, the court ruled that Section 230 immunized Google from liability for its automated editorial actions, even if those actions resulted in allegedly defamatory search snippets and dismissed the complaint with prejudice).

64 *Id*.

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

66 Ryan, *supra* note 13.

67 *Accusearch Inc.*, 570 F.3d 1187.

68 Walters v. OpenAI, LLC, No. 23-A-04860-2 (Ga. Super. Ct. May 19, 2025).

arose after ChatGPT falsely stated that a well-known radio host had embezzled funds from a gun rights organization.[69] Walters sued OpenAI for defamation, but the Superior Court of Gwinnett County in Georgia dismissed the claim holding that no reasonable person would have taken the chatbot's statements as fact especially give OpenAI's disclaimers and the plaintiff's own immediate recognition that the chat response was false.[70] The ruling reflects a growing judicial tendency to treat generative AI tools as neutral instruments rather than publishers which shifts responsibility to users to verify outputs.

Here's the crucial takeaway: while that approach may be workable for now, it leaves significant questions unanswered as AI, still in its infancy, rapidly advances and is increasingly relied upon for factual reporting and decision making. Equally notable, OpenAI had the resources to mount an extensive defense against a multi-pronged motion, but most early-stage companies do not.

The stakes become even clearer when looking outside the courtroom. In January 2025, Las Vegas police reported that a plan, originally generated by ChatGPT, to blow up a Cybertruck had been carried out.[71] Though not the subject of litigation, the incident underscores the potential for real-world harm when AI outputs go unchecked. Together, *Walters* and this extreme example show that while the law is beginning to address liability in the context of generative AI, the most pressing accountability questions remain unanswered as the technology develops at warp-speed.

*Walters* and the cases elaborated on in the aforementioned sections underscore the broad and often intricate scope of Section 230.[72] If reforms narrow Section 230's liability protections, tech companies, especially resource-constrained startups, would face greater legal exposure and be forced to invest heavily in content moderation, legal defense, and compliance systems. These added burdens could slow innovation and limit growth across the tech sector. The rise of generative AI technologies, such as Microsoft's Copilot and Google's Gemini, further complicate the landscape and push against the fabric of traditional legal frameworks. As these tools blur conventional boundaries of content creation and responsibility, they have drawn increased Congressional

---

69 *Id*.

70 *Id*.

71 Aliza Chasan, *Tesla Cybertruck Bomber Used ChatGPT to Plan Las Vegas Attack, Police Say*, CBS NEWS, (Jan. 7, 2025, 10:06 PM), https://perma.cc/GMW7-8QEP.

72 Walters v. OpenAI, LLC, No. 23-A-04860-2 (Ga. Super. Ct. May 19, 2025).

attention and fueled calls to regulate AI-generated outputs.[73] In addition, the applicability of Section 230[74] to AI-generated content is becoming a focal point in discussions about the future of internet regulation and liability frameworks.[75]

This rapid integration[76] is accompanied by escalating concerns over accountability and regulation across the globe.[77] Litigation and legislative proposals are beginning to test these boundaries. Generative AI's ability to produce realistic yet fabricated content, such as deepfakes or misleading text, has amplified calls for stricter oversight. Courts and lawmakers are grappling with the application of outdated legal frameworks to these cutting-edge technologies. Recent discussions surrounding Section 230 immunity highlight the challenges of determining whether generative AI outputs fall within the purview of "publisher" or "information content provider" protections.[78]

Furthermore, courts encounter cases that question whether the creators of generative AI systems should be held accountable for defamatory or otherwise harmful content produced by their models.[79] Legal scholars actively debate whether these systems, which rely on pre-existing datasets, merely replicate information or contribute meaningfully to the "development" of new content—a key distinction under Section 230 jurisprudence.[80] This issue remains unresolved with significant implications for developers and operators of generative AI systems.

## III. PROPOSED REFORMS

As generative AI technologies become more embedded in online platforms, courts remain skeptical of extending Section

---

73 Legislative Proposal to Sunset Section 230 of the Communications Decency Act: Hearing Before the H. Comm. on Energy & Com., 118th Cong. (May 22, 2024); *see* VALERIE C. BRANNON & ERIC N. HOLMES, CONG. RSCH. SERV., R46751, SECTION 230: AN OVERVIEW (2024); *see also* Cathy McMorris Rodgers & Frank Pallone, Jr., *Sunset of Section 230 Would Force Big Tech's Hand*, WALL ST. J. (May 13, 2024), https://perma.cc/HF7Z-SUBJ; Act of May 17, 2024, ch. 198, 2024 Colo. Sess. Laws 1199 (codified at COLO. REV. STAT. §§ 6-1-1701–1708 (2024)) (effective Feb. 1, 2026).

74 47 U.S.C. § 230.

75 *Id*.

76 *The Elastic Generative AI Report*, ELASTIC, *supra* note 9.

77 Adi Robertson, *ChatGPT Returns to Italy After Ban*, THE VERGE (Apr. 28, 2023), https://perma.cc/33XE-JTH7.

78 PETER J. BENSON & VALERIE C. BRANNON, CONG. RSCH. SERV., LSB11097, SECTION 230 IMMUNITY AND GENERATIVE ARTIFICIAL INTELLIGENCE (2023).

79 Walters v. OpenAI, LLC, No. 23-A-04860-2 (Ga. Super. Ct. May 19, 2025).

80 Ryan, supra note 13.; *see also* Louis Shaheen, *Section 230's Immunity for Generative Artificial Intelligence*, OHIO STATE UNIV., (Dec. 15, 2023).

230 immunity to AI-generated content,[81] and policymakers are actively debating legal reforms to address the unique challenges posed by AI with discussions focusing on liability, transparency, and accountability.[82] One of the most significant developments in this debate is the sunset provision proposed in the 118th Congress, which set a timeline for phasing out Section 230 protections and replacing them with new, AI-specific regulations.[83] Although this proposal was not enacted, it underscores the increasing urgency to modernize liability frameworks to reflect today's digital landscape.

### A. Establishing Clear Liability Standards for AI Platforms

A central component of the proposed legal framework is clarifying liability for platforms that use AI to generate or amplify harmful content. Section 230 currently shields platforms from responsibility for user-generated material, but certain legislators argue platforms that actively employ AI algorithms to create or promote content cannot credibly claim neutrality. For example, Rep. Frank Pallone (D-NJ), emphasized that the growing integration of generative artificial intelligence technologies into platforms will exacerbate harms and redefine the concept of a publisher, potentially introducing significant new legal challenges for companies.[84] Proposed reforms seek to hold such platforms accountable by making them liable for the outputs of their AI systems. This approach aligns with legal precedents such as *Fair Housing Council v. Roommates.com*,[85] where courts ruled that platforms that materially contribute to harmful content are not entitled to Section 230 immunity. By explicitly defining AI-driven liability, policymakers aim to ensure that platforms take ethical and safety considerations seriously in their deployment of AI technologies.

### B. Enhancing Transparency in AI Systems & State-Level Initiatives

Transparency is essential for ensuring accountability in AI governance. Lawmakers have proposed mandatory disclosures

---

81 *Walters*, No. 23-A-04860-2 (Ga. Super. Ct. May 19, 2025).

82 Legislative Proposal to Sunset Section 230 of the Communications Decency Act: Hearing Before the H. Comm. on Energy & Com., 118th Cong. (May 22, 2024).

83 *Id*.

84 *Id*. Statement of Rep. Frank Pallone.

85 *See* Fair Hous. Council of San Fernando Valley v. Roommates.com, LLC, 521 F.3d 1157 (9th Cir. 2008).

requiring platforms to inform users when AI is involved in content creation, recommendation systems, and decision-making processes. For example, California's AI Disclosure Act mandates that companies explicitly disclose AI-generated content to users with the purpose of providing greater clarity and helping consumers make informed decisions. [86]

Other state-level initiatives, such as the Colorado Artificial Intelligence Act (CAIA),[87] demonstrate the importance of localized efforts to regulate AI. Signed into law in May 2024 and effective February 2026, the CAIA is one of the first U.S. laws to directly regulate predictive AI: systems that forecast, classify, or score outcomes, such as those used in credit scoring, hiring, insurance underwriting, and public benefits eligibility. It imposes privacy, data security, and transparency safeguards while also setting specific obligations for both developers and deployers to ensure responsible design, testing, and use of AI.[88] By focusing heavily on predictive AI's role in high-stakes decision-making, the CAIA offers a regulatory model that, while distinct from generative AI governance, provides a useful precedent for assigning role-based responsibilities.[89] This predictive AI approach can inform the proportional liability framework proposed later in this Comment, which adapts similar developer-deployer distinctions but applies them to the unique risks of generative AI.

### C. Federal Reform Efforts: Section 230 Sunset Provision

As Congress explores regulatory approaches to AI governance, the proposal to sunset Section 230 by 2025, although not enacted, underscores a broader federal effort to modernize legal frameworks. This initiative aims to ensure accountability for platforms that amplify or generate harmful content through AI-driven algorithms while balancing innovation and free expression. The proposal to sunset Section 230 by 2025 was discussed extensively during the May 22, 2024, House Energy and Commerce Committee hearing.[90] Key testimony from experts and lawmakers underscored the urgent need for reform while revealing

---

86 California AI Transparency Act, S.B. 942, 2023–2024 Reg. Sess., ch. 291 (Cal. 2024) (codified at Cal. Bus. & Prof. Code § 22757).

87 Act of May 17, 2024, ch. 198, 2024 Colo. Sess. Laws 1199 (codified at COLO. REV. STAT. §§ 6-1-1701–1708 (2024)) (effective Feb. 1, 2026).

88 *Id*.

89 Act of May 17, 2024, ch. 198, 2024 Colo. Sess. Laws 1199 (codified at COLO. REV. STAT. §§ 6-1-1701–1708 (2024)) (effective Feb. 1, 2026).

90 Legislative Proposal to Sunset Section 230 of the Communications Decency Act: Hearing Before the H. Comm. on Energy & Com., 118th Cong. (May 22, 2024).

divergent perspectives on how best to balance innovation, accountability, and free expression in the generative AI era. Chair Cathy McMorris Rodgers (R-WA) emphasized the bipartisan intent behind sunsetting Section 230: creating a definitive timeline to ensure meaningful reforms. She noted the growing harms caused by the unchecked amplification of harmful content, particularly through AI-powered algorithms.[91] "Big tech has failed to uphold American values," Rep. Rodgers stated, " . . . and they must be good stewards of their platforms."[92] Ranking Member Frank Pallone (D-NJ) added that the sunset proposal is not about erasing Section 230 but ensuring Congress develops a framework that reflects today's internet realities.[93] Witness Marc Berkman supported the sunset strategy, and he described it as a "public health necessity" given the pervasive harms exacerbated by tech platforms.[94] Without reform, he warned, the current imbalance favoring profit over public safety will persist.

Recognizing that traditional legal frameworks may not adequately address the unique challenges posed by generative AI, some legislators have called for tailored regulatory standards. Rep. Darren Soto (D-FL) proposed creating distinct legal obligations for platforms actively generating or amplifying content and distinguished them from those merely hosting third-party material.[95] This distinction allows the law to account for the specific risks associated with AI-driven platforms while preserving protections for smaller entities or passive intermediaries. A tailored approach ensures that regulatory measures address the realities of AI systems without imposing unnecessary burdens on businesses that operate within traditional parameters.

As previously mentioned, the proposal to sunset Section 230 was not enacted, but it exemplifies congressional interest in broader platform liability reforms. Lawmakers continue to explore alternative measures to address the challenges posed by AI-driven content amplification. Several legislative proposals, including those targeting algorithmic recommendations, reflect ongoing efforts to refine liability standards and adapt existing legal frameworks to the evolving digital landscape. These initiatives highlight the continued search for solutions that balance innovation, accountability, and free expression.

---

91 *Id*. Statement of Rep. Cathy McMorris Rodgers.

92 *Id*.

93 *Id*. Statement of Rep. Frank Pallone.

94 *Id*. Statement of Marc Berkman, CEO & Founder, Org. for Soc. Media Safety.

95 *Id*. Statement of Rep. Darren Soto.

D. Beyond the Sunset Provision: Other Federal Reform Efforts

Lawmakers in both the 117th and 118th Congresses have pursued legislative efforts to limit Section 230 protections for platforms that algorithmically curate or amplify content. These proposals, including the reintroduced DISCOURSE Act,[96] seek to narrow immunity for online services that actively recommend or promote certain material; these proposals reflect growing concerns over the role of algorithms in content dissemination. Table 1 focuses on legislative proposals from the 117th Congress aimed at algorithmic recommendations under Section 230. It does not include broader reforms or bills addressing content moderation and access restrictions.

**Table 1. Proposed Legislation Addressing Section 230 from the 117th Congress & Algorithmic Recommendations**[97]

| Bill No. | Short Title | Function |
|---|---|---|
| **H.R. 9695** | Platform Integrity Act | Proposed changes to Section 230(c)(1) aimed at restricting legal protections when a service provider or user has "promoted, suggested, amplified, or otherwise recommended" the disputed content. This would have narrowed the scope of liability immunity for platforms actively engaging with content distribution. |
| **H.R. 5596** | Justice Against Malicious Algorithms Act | Sought to limit Section 230(c)(1) protections for platforms that knowingly or recklessly made a "personalized recommendation" that played a significant role in causing "physical or severe emotional injury to any person." However, this restriction would not have applied to cases where recommendations resulted "directly in response to a user-specified search." |

96 S. 2228, 117th Cong. (2021); reintroduced as S. 921, 118th Cong. (2023).

97 HOLMES, *supra* note 4.

| | | |
|---|---|---|
| **S. 2335** | Don't Push my Buttons Act | Proposed that platforms lose immunity when a provider "(i) collects information regarding the habits, preferences, or beliefs of a user of the service; and (ii) uses an automated function to deliver content to the user described in clause (i) that corresponds with the habits, preferences, or beliefs identified as a result of the action taken under that clause with respect to that user." However, the exemption would not apply if a user "uses an automated function to deliver content to that user" or "knowingly and intentionally elects to receive the content." |
| **S. 2228 Reintroduced in 118th Congress as S.921** | DISCOURSE Act | Included a provision to redefine "information content provider" to encompass a platform "with a dominant market share" that employs certain algorithms to target content. This legislation was later reintroduced in the 118th Congress as S. 921. |
| **H.R. 2154** | Protecting Americans from Dangerous Algorithms Act | Intended to revoke liability protections for interactive computer services in certain federal civil lawsuits involving civil rights violations and terrorism when platforms utilized algorithms to sort, prioritize, or recommend third-party content. The bill contained specific carve-outs and exemptions. |
| **H.R. 2448** | Health Misinformation Act | Aimed at holding platforms accountable if they "promote . . . health misinformation through an algorithm" during an officially declared public health emergency. This measure sought to curb the spread of misleading health-related information by removing immunity under Section 230 in such contexts. |

As Table 1 illustrates, several legislative initiatives from the 117th Congress aimed to narrow Section 230's protections by targeting how platforms recommend, rank, or surface content. While the proposals varied in phrasing, they shared a common concern: that algorithmic systems can influence visibility and user experience in ways that resemble editorial judgment. However, commentators have noted the difficulty of determining what counts as amplification given the technical differences across content recommendation.[98] Critics of reform argue that simply organizing or displaying content on a platform naturally boosts or amplifies some posts over others simply as a product of platform design.[99] For example, if a user opens Reddit, the algorithm determines which posts appear at the top of their feed or subreddit. Even if Reddit is not actively "promoting" a post, its ranking system naturally makes some posts more visible than others based on factors like upvotes, comment activity, post age, and user preferences.[100] Judicial rulings in *Force*[101] and *Gonzalez*[102] adopted a similar perspective concluding that algorithmic curation constitutes a form of editorial discretion protected under Section 230(c)(1).[103] A key issue remains whether exemptions from Section 230 would extend to search engines, which are inherently designed to prioritize and surface results deemed most relevant to a user's query.

This debate over algorithmic curation highlights a broader regulatory tension: most of these legislative proposals implicitly address predictive AI—systems that forecast user preferences and deliver ranked or recommended content—rather than generative AI, which creates entirely new content such as text, images,

---

98 *See* Daphne Keller, *Amplification and Its Discontents*, 1 J. FREE SPEECH L. 227, 232–33 (2021); *see also* Manoel Horta Ribeiro, Veniamin Veselovsky & Robert West, *The Amplification Paradox in Recommender Systems* (ar Xiv: 2302.11225 [cs.CY]) (contending that interpretations of algorithmic amplification should consider user engagement with suggested content and asserting that recommendation systems are "not the main factor directing attention to extreme content").

99 Eric Goldman, *Search Engine Bias and the Demise of Search Engine Utopianism*, 8 YALE J.L. & TECH. 188, 195–96 (2006) (arguing "search engines simply cannot passively and neutrally redistribute third party content" and explaining search engines, like Google, rank results based on algorithms that prioritize certain websites over others. Even if they are not deliberately promoting content, their design inherently favors some links by placing them at the top, influencing what users see first and shaping perceptions of relevance).

100 Alex Moehring, *Personalization, Engagement, and Content Quality on Social Media: An Evaluation of Reddit's News Feed*, MIT SLOAN SCH. OF MGMT. (Working Paper, May 30, 2024).

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

102 Gonzalez v. Google LLC, 2 F.4th 871 (9th Cir. 2021).

103 47 U.S.C. § 230 (c)(1).

or audio. Predictive AI raises concerns about bias, filter bubbles, and targeted harm through recommendation, whereas generative AI introduces distinct challenges around authorship, misinformation creation, and synthetic media. This distinction is critical for shaping a proportional liability framework since the regulatory levers that work for predictive models (e.g., ranking transparency, recommendation limits) may be ineffective or incomplete for governing generative systems that originate new expressive works. The current legal framework, built for human-created and user-submitted material, remains ill-equipped to address the distinct risks posed by content produced directly by non-human actors since existing laws were primarily designed to address user-generated content, not AI-generated speech, deepfakes, or autonomous decision-making systems. This fundamental shift underscores the need for updated legislation that reflects the distinct challenges posed by generative AI, particularly regarding liability, transparency, and accountability.

Regulators worldwide are increasingly recognizing this challenge. As illustrated in the United States, various state-level efforts have proposed measures to enhance transparency, fairness, and accountability in AI systems.[104] Internationally, the European Union's Artificial Intelligence Act (Regulation (EU) 2024/1689) takes a risk-based approach to AI governance by categorizing AI systems based on their potential harm and imposing strict obligations on high-risk applications.[105] Generative AI, given its capacity to produce deceptive, biased, or harmful outputs at scale, is likely to fall into this high-risk category.[106] The current legal framework in the U.S. appears ill-suited to address the realities of content generated by non-human actors. This underscores the urgent need for updated legislation that accounts for the distinct risks posed by generative AI, particularly in determining liability structures and the appropriate regulatory safeguards.

While concerns over limiting Section 230 protections persist, reform advocates emphasize that these proposals address significant gaps in platform accountability. Supporters argue that

---

[104] California AI Transparency Act, S.B. 942, 2023–2024 Reg. Sess., ch. 291 (Cal. 2024) (codified at Cal. Bus. & Prof. Code § 22757).

[105] Regulation 2024/1689, of the European Parliament and of the Council, Laying Down Harmonized Rules on Artificial Intelligence and Amending Regulations (EC) No. 300/2008, (EU) No. 167/2013, (EU) No. 168/2013, (EU) 2018/858, (EU) 2018/1139, and (EU) 2019/2144, and Directives 2014/90/EU, (EU) 2016/797, and (EU) 2020/1828, 2024 O.J. (L 1689) 1 (July 12, 2024).

[106] *Id*.

narrowing immunity for algorithmically amplified content strengthens consumer protections, deters the spread of harmful misinformation, and incentivizes platforms to adopt safer AI moderation practices.[107] Proposals such as the Justice Against Malicious Algorithms Act reflect growing bipartisan agreement that platforms should not remain entirely shielded from liability when their AI systems contribute to real-world harm.[108] Moreover, regulatory supporters contend that adjusting liability for content amplification fosters a more competitive market by preventing dominant platforms from exploiting algorithmic advantages to the detriment of smaller competitors. However, critics highlight the unintended consequences of these reforms including increased litigation risks for startups, compliance burdens that may stifle innovation, and potential over-removal of lawful content due to liability concerns.[109] These trade-offs demonstrate the complex challenge of balancing platform responsibility with the need to maintain a vibrant digital ecosystem.

At the same time, emerging technologies like AI-generated content introduce new dimensions to the liability debate which further complicate existing legal frameworks. While Section 230 was originally designed to govern third-party speech, its applicability to AI-generated material remains uncertain. As lawmakers grapple with content moderation reforms, the rise of AI-driven platforms raises pressing questions about accountability in an era where content creation is increasingly automated. This shift underscores the need for updated liability structures that reflect the evolving role of AI while preserving fundamental protections for innovation and online expression. Lawmakers now face the challenge of determining whether a proportional liability model, similar to the EU's risk-based AI regulation, provides a viable solution to these emerging concerns.[110]

These legal uncertainties are particularly relevant for industries that rely heavily on Section 230 protections. The potential narrowing of these immunities could disrupt key business models and change how platforms manage content and interact with users. Companies that depend on user-generated content, recommendation algorithms, and AI-driven moderation face heightened

---

[107] Legislative Proposal to Sunset Section 230 of the Communications Decency Act: Hearing Before the H. Comm. on Energy & Com., 118th Cong. (May 22, 2024) (statement of Marc Berkman, CEO & Founder, Org. for Soc. Media Safety).

[108] Justice Against Malicious Algorithms Act of 2021, H.R. 5596, 117th Cong. (2021).

[109] Huddleston, *supra* note 45; *see also* Nabil, *supra* note 45; Doctorow, *supra* note 45.

[110] CITI Program Staff, *An Overview of the EU AI Act: What You Need to Know*, CITI PROGRAM (Jan. 2, 2025), https://perma.cc/57ZQ-VSEY (last visited Jan. 18, 2025).

risks as legal responsibilities shift. Without clear guidelines, platforms may need to overhaul their operational structures, invest heavily in compliance mechanisms, or limit certain features to avoid liability. As discussions around reform continue, it is crucial to examine which business models stand to be most affected by changes to Section 230. From aggregators and forums to social media platforms and AI content generators, each relies on these protections to function efficiently. Understanding how these industries navigate liability concerns provides insight into the broader implications of modifying Section 230.

## IV. BUSINESS MODELS WHERE SECTION 230 MATTERS MOST

Business models that heavily rely on Section 230 protections include aggregators and forums, social media platforms, and creators of large LLMs, all of which depend on immunity to operate effectively. Aggregators and forums like Reddit and Craigslist use these protections to maintain open user discussions without the financial burden of extensive content moderation. Social media giants such as Meta and YouTube rely on Section 230 to shield their content recommendation algorithms from liability, which allows them to prioritize user engagement and growth. Additionally, LLM creators like OpenAI and Anthropic navigate the complexities of AI-generated content since they operate in a gray area where their outputs blur the lines between user-generated and platform-produced content.[111] Proposed reforms to Section 230 threaten to disrupt these business models by increasing operational costs and legal risks while stifling innovation in these industries.

### A. Aggregators and Forums

Aggregators and forums, such as Reddit and Craigslist, are quintessential examples of platforms thriving on user-generated content to maintain their engagement and relevance. These platforms function as digital bulletin boards: hosting a diverse range of content with minimal editorial oversight. Section 230 has historically shielded these companies from liability for the vast majority of user-generated posts, so they can operate without the prohibitive costs of extensive content moderation.[112]

However, reforms to Section 230 would fundamentally alter this business model. If aggregators and forums are held liable for

---

111 OpenAI et al., *supra* note 30.

112 *See generally* Zeran v. Am. Online, Inc., 129 F.3d 327 (4th Cir. 1997).

user-generated content, they would face significant operational burdens including the need to implement robust content monitoring systems. For smaller platforms, the cost of deploying AI-powered moderation tools or employing human moderators could outweigh the benefits of maintaining open forums. Moreover, these platforms might be forced to restrict user interactions to mitigate liability, thereby reducing the value they provide to their communities.

### B. Content Producers

Platforms such as YouTube and Meta operate as major content producers in the digital ecosystem, even though most of the material comes from their users. Their business models depend on sophisticated recommendation systems, often powered by AI, that surface and promote content to keep people engaged. Section 230 has been central to this model since it protects them from liability when those algorithms elevate harmful or defamatory material.[113]

Proposed reforms to Section 230, including carveouts for algorithm-driven promotion, pose existential risks to these platforms. If content producers are held accountable for the outcomes of their recommendation systems, they may need to invest heavily in algorithmic transparency and redesign their systems to avoid promoting harmful content. This shift could significantly increase operational costs and reduce the efficiency of their algorithms, ultimately impacting user satisfaction and engagement. For example, if Meta were required to disclose the inner workings of its content recommendation system, it could face competitive disadvantages and challenges in safeguarding proprietary technology. At the same time, YouTube might be forced to demote user-uploaded content to ensure compliance with new liability standards which may curb the platform's ability to support content creators.

### C. Large Language Model Creators

Generative AI companies like OpenAI and Google DeepMind sit at a critical juncture in the debate over Section 230 protections

---

[113] Gonzalez v. Google LLC, 2 F.4th 871 (9th Cir. 2021) (Recall, this case directly addressed Google's liability for its content recommendation algorithms under Section 230. The plaintiffs argued that Google's algorithm promoted ISIS-related content, contributing to a terrorist attack. The court ultimately held that Section 230 provided immunity for Google's algorithmic recommendations, underscoring the pivotal role of Section 230 in protecting platforms like YouTube and Meta from liability tied to their content curation and promotion practices.).

and defamation law.[114] Unlike traditional platforms, these companies create tools that actively produce content based on user prompts resulting in novel legal issues around liability for AI-generated content.[115] For example, defamation law traditionally requires a false and harmful statement presented as fact, published to a third party, with some level of fault on the publisher's part.[116] Applying these standards to generative AI outputs, or "hallucinations," involves complex questions such as whether AI-generated outputs can be considered factual statements particularly when companies like OpenAI use disclaimers to emphasize the probabilistic nature of their models.[117]

Cases like *Walters v. OpenAI* and *Battle v. Microsoft* illustrate the emerging challenges in this area.[118] In *Walters*, OpenAI argued that ChatGPT functions as a private drafting tool rather than a publisher and relied on terms of use that assign responsibility to the user.[119] Meanwhile, in *Battle*, the plaintiff alleged reputational harm caused by Bing's AI-assisted search engine, which conflated his biography with that of another individual.[120] Although the *Battle* case was stayed for arbitration, it highlights the potential reputational risks generative AI outputs can create and the demand for remedies like permanent removal of false information.[121]

If courts mandate that AI companies adopt notice-and-takedown systems or enhance content moderation safeguards, the resulting operational costs could become a significant burden especially for smaller companies with limited resources. The challenge lies in balancing the need for accountability with fostering innovation as courts and lawmakers attempt to apply defamation law to generative AI tools. Central to this issue is how courts interpret key legal concepts such as "publication" and fault as well as how users perceive the factual reliability of AI-generated content in various contexts. To mitigate potential liability, companies developing LLMs might need to incorporate advanced safeguards like real-time content filtering, bias detection, and robust

---

114 Walters v. OpenAI, LLC, No. 23-A-04860-2 (Ga. Super. Ct. May 19, 2025).

115 *See* Battle v. Microsoft, No. 1:23-cv-01822 (D. Md. filed July 7, 2023).

116 Restatement (Second) of Torts § 558 (Am. L. Inst. 1977).

117 *See* OpenAI, Terms of Use, https://perma.cc/ZR4X-LXHU (last visited Jan. 3, 2025).

118 *Walters*, No. 23-A-04860-2 (Ga. Super. Ct. May 19, 2025); *Battle*, No. 1:23-cv-01822 (D. Md. filed July 7, 2023).

119 *Walters*, No. 23-A-04860-2 (Ga. Super. Ct. May 19, 2025).

120 *Battle*, No. 1:23-cv-01822 (D. Md. filed July 7, 2023).

121 *Id*.

monitoring systems. Although such measures could help address legal risks, they would also substantially increase development expenses and slow the release of new AI technologies. For smaller startups in the generative AI space, these heightened requirements could create insurmountable barriers and force them to compete on uneven footing with larger, well-funded corporations and potentially stifle tech diversity and innovation in the field.

## V. IMPLICATIONS FOR STARTUPS

### A. Increased Liability and Operational Costs

The operational burden and litigation costs associated with intermediary liability present significant challenges for startups with pre-complaint costs ranging from $0 to $3,000, motion-to-dismiss expenses between $15,000 and $80,000, early motions for summary judgment costing $30,000 to $150,000, and discovery through trial reaching $100,000 to over $500,000, which makes litigation an increasingly costly "lose-lose" scenario that often forces settlements even in weak cases.[122] These expenses can strain early-stage budgets, divert resources from growth, and create strong incentives to resolve disputes preemptively even when claims lack merit. Without Section 230 protections, startups would also face increased operational burdens, including the need for extensive content moderation and legal compliance programs, which could siphon resources away from innovation and growth.

### B. Chilling Effect on Innovation

The removal or significant narrowing of Section 230 protections risks stifling innovation and disrupting market dynamics. According to Engstrom, Section 230 protects startups from being overwhelmed by frivolous lawsuits and allows them to allocate resources toward innovation rather than legal defense.[123] Without these protections, startups would struggle to survive in a legal environment where only well-funded firms could afford the high compliance and litigation expenses. Commentators such as Engstrom and Huddleston argue that increased liability risks consolidating power within a few dominant players, effectively pushing

---

[122] *See* Engstrom, *supra* note 42 (breaking down the cost of litigation for startup and early-stage companies).

[123] Fair Hous. Council of San Fernando Valley v. Roommates.com, LLC, 521 F.3d 1157, 1174 (9th Cir. 2008).

smaller, resource-constrained companies out of the market.[124] This reduction in competition would hinder diversity in the digital economy, lead to less innovation, and reduce consumer choices.

### C. The Legal Burden on Startups in the AI Era

Proposed AI regulations further exacerbate the challenges startups face and create compliance requirements that favor larger, well-resourced companies. While legislators aim to balance fostering innovation with protecting the public from AI-related harms, the current proposals disproportionately burden smaller firms. These frameworks introduce complex legal standards that startups often lack the financial and operational capacity to implement effectively. Key regulatory measures include liability clarification, transparency mandates, algorithmic oversight, and tailored legal standards for AI platforms. While these efforts seek to create guardrails for responsible AI deployment, they inadvertently place an outsized burden on startups, which lack the legal infrastructure to manage such obligations.

### D. Addressing Current Regulatory Frameworks

Regulations designed to promote fairness and accountability in AI governance create barriers that disproportionately affect startups. Many proposals impose extensive compliance requirements that favor larger corporations with the legal and financial resources to navigate regulatory complexities.[125] Without tailored provisions, new laws could reinforce market concentration. Regulatory frameworks need to incorporate scaled compliance obligations based on company capacity, phased rollouts, and good faith provisions that ensure accountability without stifling innovation.

Proponents of strict AI regulations argue that these measures protect consumers, prevent misinformation, and ensure responsible AI deployment.[126] Without oversight, AI technologies could amplify bias, compromise privacy, and contribute to widespread societal harm. By creating accountability mechanisms, regulators seek to mitigate these risks while fostering a transparent and trustworthy AI ecosystem.

---

[124] Huddleston, *supra* note 45; *see also* Nabil, *supra* note 45; Doctorow, *supra* note 45.

[125] Legislative Proposal to Sunset Section 230 of the Communications Decency Act: Hearing Before the H. Comm. on Energy & Com., 118th Cong. (May 22, 2024).

[126] *Id*. Statement of Marc Berkman, CEO & Founder, Org. for Soc. Media Safety.

Consumer protection remains one of the strongest arguments for AI regulation.[127] AI-generated content has fueled disinformation campaigns, deepfake manipulation, and biased decision-making systems. Oversight aims to prevent companies from prioritizing speed and profitability over ethical responsibility. Furthermore, broad compliance requirements set an industry-wide baseline for responsible AI practices by ensuring companies address ethical risks before they become systemic problems rather than retroactively attempting to correct harms.

Regulatory fairness is another justification for AI oversight. Large technology companies already control vast amounts of data and computing resources, which strengthens their market position.[128] Without clear rules, these firms can exploit their advantages to entrench dominance and wipeout new competitors. Enforcing transparency and accountability across all AI companies can reduce these disparities by requiring every organization, regardless of capacity, to adopt responsible practices.

Although oversight aims to prevent abuse, strict regulatory mandates introduce unintended consequences. Compliance costs and legal uncertainties create significant obstacles for startups and limit their ability to innovate and compete. Regulatory frameworks that fail to adjust requirements based on company capacity and risk exposure risk consolidating power among a few dominant firms. This concern is particularly acute in the AI industry where market power is already concentrated. A small number of companies, such as OpenAI, Google DeepMind, Anthropic, and Meta, dominate access to foundational models, while upstream dependencies on chipmakers like NVIDIA further entrench concentration in the supply chain.[129] Regulations structured around scalable compliance obligations offer a potential solution. Startups with limited financial and operational capacity should receive extended compliance timelines, reduced reporting requirements, and access to regulatory guidance. By tailoring obligations to match a company's capacity and resources, regulators can maintain high ethical standards while ensuring smaller firms can participate in AI development.

A phased regulatory rollout provides another strategy to prevent unintended market consolidation. Gradually implementing

---

[127] *Id*.

[128] ZHU & HARRIS, *supra* note 49.

[129] FED. TRADE COMM'N, OFFICE OF TECH. STAFF, Partnerships Between Cloud Service Providers and AI Developers: 6(b) Study Staff Report (Jan. 2025), https://perma.cc/U4HB-4BQH.

compliance requirements allows startups to adapt, integrate best practices, and innovate while avoiding sudden regulatory shocks. This approach enables policymakers to refine rules based on industry feedback and technological advancements rather than imposing rigid standards that may become outdated. Good faith provisions also provide flexibility for startups. Companies that demonstrate efforts to mitigate risks should not face immediate legal consequences if they struggle to meet full compliance requirements. Recognizing ethical AI development rather than imposing strict liability ensures that startups remain competitive without undermining consumer protections. These provisions hold bad actors accountable while protecting early-stage companies from undue regulatory pressure.

AI regulation must balance accountability and innovation. While oversight is essential for ethical AI deployment, overly rigid rules risk consolidating industry control within the largest tech firms. A regulatory environment that incorporates proportional compliance measures, phased implementation, and good faith protections encourages responsible AI development while fostering competition. Without these adjustments, AI governance risks reinforcing monopolistic control, limiting market diversity, and slowing technological advancement.

## VI. A PROPORTIONAL REGULATORY FRAMEWORK: BALANCING AI RISK, COMPANY CAPACITY, AND SANDBOX INNOVATION

The rapid spread of artificial intelligence has forced regulators to confront a difficult question: how do we create rules that keep the public safe without hindering innovation? The answer isn't one-size-fits-all because not all AI works the same way. Generative AI can produce brand new material from songs to convincing deepfakes in a matter of seconds. On the other hand, predictive AI uses data to forecast outcomes or make recommendations, like deciding which ads you see online, estimating your creditworthiness, or flagging suspicious activity. Both raise risks, but the problems they create and the safeguards they require are not identical.

Some U.S. states have started to recognize this distinction. CAIA, for example, is aimed primarily at predictive AI and seeks to protect people from harmful outcomes by requiring developers and deployers in areas like housing, employment, and lending to be transparent, test their systems for bias, and give individuals notice and recourse when AI influences important decisions such

as jobs, loans, or healthcare.[130] The same approach could apply to generative AI, but generative systems raise different challenges, like tracing the source of outputs, auditing model behavior, and guarding against prompt manipulation, that go beyond CAIA's current focus.[131]

At the same time, federal lawmakers have been wrestling with whether states should even be permitted to move ahead with their own AI regulations.[132] Earlier this year, Congress debated a proposal that would have placed a ten-year freeze on new state-level AI laws.[133] Supporters framed the moratorium as a way to prevent a patchwork of conflicting rules that could slow innovation and create compliance headaches for companies operating across state lines.[134] The idea, however, quickly ran into resistance.[135] Critics warned that such a sweeping pause would come at too high a cost by stripping states of their ability to protect their residents at a time when AI tools are rapidly entering sensitive areas of daily life.[136] They pointed to existing state privacy laws, online safety protections for children, and consumer rights statutes as examples of rules that could be weakened or preempted under a federal moratorium.[137]

The Senate eventually voted to strip the moratorium language from a major budget bill[138] with opponents from both parties emphasizing that states have long played a vital role as early laboratories of regulation.[139] The collapse of the moratorium effort underscores just how divided federal policymakers remain: some see state regulation as a threat to innovation while others view it as a necessary safeguard in the absence of comprehensive federal rules. For now, states like Colorado are leading the way, but the debate reflects a deeper tension between the desire for a unified

---

130 Act of May 17, 2024, ch. 198, 2024 Colo. Sess. Laws 1199 (codified at COLO. REV. STAT. §§ 6-1-1701–1708 (2024)) (effective Feb. 1, 2026).

131 *Id*.

132 Roll Call Vote on S. Amdt. 2814 to S. Amdt. 2360 to H.R. 1, 119th Cong., 1st Sess. (July 1, 2025), https://perma.cc/N4WR-WJY9.

133 *Id*.

134 U.S. SENATE COMM. ON COMMERCE, SCI. & TRANSP., Press Release, *Sen. Cruz: Adopting Europe's Approach on Regulation Will Cause China to Win the AI Race* (May 8, 2025), https://perma.cc/EBC2-JWE8.

135 U.S. SENATE COMM. ON COMMERCE, SCI. & TRANSP., Press Release, *Ranking Member Cantwell Says Blackburn-Cruz AI Moratorium Amendment Does Nothing to Protect Kids and Consumers* (June 30, 2025), https://perma.cc/2Q8N-GAYX.

136 *Id*.

137 *Id*.

138 *See* One Big Beautiful Bill Act, H.R. 1, 118th Cong. (2025).

139 Roll Call Vote on S. Amdt. 2814 to S. Amdt. 2360 to H.R. 1, 119th Cong., 1st Sess. (July 1, 2025), https://perma.cc/N4WR-WJY9.

federal framework and the value of state-by-state experimentation in shaping AI governance. Across the Atlantic, lawmakers have moved in a different direction.

The European Union's AI Act establishes a single, risk-based framework that applies uniformly across member states.[140] Rather than distinguishing between predictive and generative systems, the Act sorts all AI into four risk tiers: unacceptable, high, limited, and minimal.[141] Higher-risk systems face stricter obligations, including human oversight, accuracy testing, and post-market monitoring.[142] While the Act does include some small and medium-sized enterprise (SME) accommodations, such as regulatory sandboxes and reduced documentation, these remain peripheral to the core risk-tiering approach.[143] The result is that a large multinational corporation and a small startup face essentially identical compliance obligations when they deploy the same high-risk system. This approach may work for big firms with legal and compliance teams, but it can overwhelm smaller players, discourage innovation, and tilt the market toward incumbents. Notably, the Act explicitly references startups within its SME provisions, even though EU law lacks a standalone definition of what qualifies as a "startup."

### A. Company Capacity & Sandbox Innovation

My framework builds on the EU's risk-based model but adapts it in two important ways. First, it introduces company capacity as a second organizing principle. Capacity refers to a firm's size, resources, market reach, and its role in the AI lifecycle: whether it is building models, integrating them into products, or deploying them directly to consumers. These factors shape what the organization can reasonably be expected to do to manage risk. A global technology company rolling out a high-risk system should shoulder far heavier obligations, including rigorous testing, ongoing monitoring, and strong redress mechanisms, as opposed to a small startup experimenting with the same technology. Startups would still need to meet core safeguards, but their

---

[140] Regulation 2024/1689, of the European Parliament and of the Council, Laying Down Harmonized Rules on Artificial Intelligence and Amending Regulations (EC) No. 300/2008, (EU) No. 167/2013, (EU) No. 168/2013, (EU) 2018/858, (EU) 2018/1139, and (EU) 2019/2144, and Directives 2014/90/EU, (EU) 2016/797, and (EU) 2020/1828, 2024 O.J. (L 1689) 1 (July 12, 2024).

[141] *Id*.

[142] *Id*.

[143] *Id*.

obligations would scale to what is realistically within reach. Embedding capacity alongside risk keeps the principle of proportionality at the center of enforcement.

Second, this framework gives regulatory sandboxes a more central role. Under the EU AI Act, a sandbox is defined as a controlled environment created by regulators that allows providers to develop, train, and test AI systems in real-world conditions for a limited time under close supervision.[144] These sandboxes are included as supportive measures for SMEs, but they operate at the margins of the Act's main classification system.[145] In practice, that means a startup and a multinational corporation deploying the same high-risk AI tool face nearly identical compliance requirements with sandboxes offering only a temporary reprieve.

The U.S. has taken a different path to sandboxes in another context: financial technology.[146] State-level fintech sandboxes in places like Arizona, Utah, and Wyoming were designed to give new entrants breathing room by allowing them to test financial products under regulatory oversight without being crushed by the full weight of existing laws from day one. These programs still demanded baseline consumer protections, but they created space for young firms to innovate responsibly while regulators learned alongside them. That balance, guardrails without paralysis, is instructive for AI.

My proposal merges these strands. It retains the EU's risk-based categories, but it overlays them with capacity tiers and ties them directly to sandbox participation. Larger companies would have fewer opportunities to rely on sandboxes since they have the resources to meet full compliance from the start. Smaller firms, however, could enter supervised sandboxes where they must meet core safeguards, such as bias testing, transparency notices, and redress pathways, while being given room to experiment and scale. For size determinations, the framework draws on the U.S. Small Business Administration (SBA) Size Standards.[147] The SBA

---

144 Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 on Artificial Intelligence, 2024 O.J. (L 202) 1, ch. VI, arts. 57–63 (EU) (establishing regulatory sandboxes and related innovation-support measures including special provisions for SMEs and startups).

145 *Id*.

146 *See, e.g.*, ARIZ. REV. STAT. ANN. §§ 41-5601 to -5612 (2023); FLA. STAT. § 559.952 (2023); UTAH CODE ANN. §§ 13-55-101 to -108 (West 2023); WYO. STAT. ANN. §§ 40-28-103 to -109 (2023); NEV. REV. STAT. § 657A.100 (2023); W. VA. CODE §§ 31A-8G-1 to -8 (2023).

147 *Size Standards*, U.S. SMALL BUS. ADMIN., https://perma.cc/3JH9-V5Q3 (last visited Jan. 18, 2025).

provides a Size Standards Tool,[148] which enables businesses to determine whether they qualify as a small business based on their North American Industry Classification System (NAICS) code.[149] In brief, the sandbox approach ensures that innovation is not stifled at its earliest stages. Equally important, public trust is not sacrificed in the process.

The specific sandbox for startups would look something like this: a temporary program that lets early-stage companies test new AI tools with real users in a safe, closely monitored environment. Instead of being subject to the full weight of regulation right away, startups could operate under tailored rules for a set period, say twelve to eighteen months, while regulators watch closely, provide guidance, and ensure basic protections like fairness testing and user transparency. At the end, the company would receive feedback and an official report to help it decide whether to scale, pivot, or seek the standard regulatory approvals needed to operate permanently. This way, young firms get the breathing room to experiment while the public still benefits from guardrails that prevent harmful outcomes.

### B. Good Faith Provision

A central feature of this framework is the good faith provision, which ensures that AI deployers, regardless of their tier, take reasonable measures to prevent harm. This provision establishes a foundational principle that all companies, regardless of size or resources, must act responsibly when developing and deploying AI technologies. It serves as a safeguard against negligent or unethical AI practices while promoting fairness in regulatory obligations.[150]

Under this provision, all AI deployers are required to take reasonable precautions in the deployment of their systems.[151] Reasonable, in this context, means actively conducting risk

---

148 *Size Standards Tool*, U.S. SMALL BUS. ADMIN., https://perma.cc/94VR-L364 (last visited Jan. 18, 2025); *see also Table of Size Standards*, U.S. SMALL BUS. ADMIN., https://perma.cc/PE2J-CCR9 (last visited Jan. 18, 2025).

149 *North American Industry Classification System (NAICS)*, U.S. CENSUS BUREAU, https://perma.cc/8LP2-GN4M (last visited Jan. 18, 2025).

150 ZHU & HARRIS, *supra* note 49.

151 Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024, laying down harmonized rules on artificial intelligence and amending various Regulations and Directives, 2024 O.J. (L 1689) 35, ¶ 140, (requiring providers in AI regulatory sandboxes to act in *good faith* by following guidance from competent authorities and promptly mitigating significant risks to safety, health, and fundamental rights during AI system development, testing, and experimentation).

assessments, implementing appropriate safeguards, establishing clear oversight mechanisms, and regularly reviewing AI performance to identify and mitigate potential risks. These measures should be well-documented, continuously updated, and scaled to match the complexity and potential impact of the AI system in question. By mandating these actions, the framework ensures that AI technologies are designed and operated in ways that prioritize user safety and societal well-being.

The good faith provision also plays a critical role in preventing bad actors from exploiting lenient requirements. Without such a safeguard, companies might attempt to circumvent regulatory obligations or manipulate AI systems in ways that could cause harm. By enforcing a standard of good faith, the framework holds all companies accountable for their AI deployments and reduces opportunities for misuse while still allowing flexibility for ethical innovation. Companies that fail to conduct risk assessments, deliberately obscure AI decision-making processes, or neglect transparency measures would not be operating in good faith and would face stricter liability consequences.

Additionally, this provision is designed to protect startups from undue liability and guarantee that they are not disproportionately penalized for unintentional harms. Given that smaller companies often have limited resources compared to larger corporations, they may face challenges in fully addressing all potential AI risks. The good faith provision ensures that if these companies make a sincere effort to adhere to best practices, follow reasonable risk mitigation procedures, and deploy AI ethically, they will not be unfairly targeted by excessive regulatory burdens or legal consequences.

Beyond liability protection, the good faith provision establishes a baseline of ethical AI practices including transparency and accountability. Companies are expected to maintain clear documentation of their AI development and decision-making processes by providing users and regulators with insight into how their AI systems operate. This includes disclosing key factors influencing AI decisions, maintaining explainability in system outputs, and allowing for human oversight where necessary. Such transparency fosters public trust in AI technologies while enabling regulators to assess compliance effectively.

Finally, the good faith provision aligns with risk scaling by ensuring that larger companies deploying high-risk AI systems must demonstrate greater due diligence than startups using low-risk AI. High-risk AI applications, such as those involved in

hiring, financial decisions, or biometric surveillance, require extensive oversight, regular auditing, and stringent compliance mechanisms. In contrast, low-risk AI applications, such as automated recommendations or spell-checking tools, may have fewer regulatory requirements. This risk-adjusted approach ensures that regulatory expectations are proportional to the potential impact of AI deployments while maintaining a commitment to ethical AI development across all tiers.

### C. Framework Enhancements

To further align my framework with best regulatory practices several improvements could be made. One enhancement involves refining the prohibited AI uses category. Similar to the EU AI Act's unacceptable risk tier, this refinement would provide clear guidance on which AI applications should be outright banned due to their potential for significant societal harm.[152] This addition would help clarify the boundaries of ethical AI deployment while reinforcing accountability for high-risk applications.

Another improvement would be further aligning liability tiers with SBA standards.[153] Explicitly linking liability obligations to SBA definitions would provide clearer regulatory expectations for businesses of different capacities. This adjustment would ensure that compliance requirements remain fair and achievable while maintaining regulatory oversight where needed. Introducing compliance incentives could also strengthen the framework. Startups that prioritize ethical AI practices could benefit from regulatory safe harbors, grants, or phased compliance measures. By rewarding companies that take proactive steps to ensure AI safety and transparency, this enhancement would create a positive incentive structure that promotes responsible AI deployment.

## VII. CONCLUSION

Generative AI is forcing courts and lawmakers to confront questions that Section 230 was never built to answer. Treating AI outputs as if they were ordinary user posts leaves major risks unchecked, but tearing down immunity altogether would shut the door on smaller firms that cannot survive the costs of constant legal exposure. Both approaches miss the middle ground.

---

[152] *Id*.

[153] *Size Standards*, *supra* note 147; see also *Size Standards Tool*, *supra* note 148; *Table of Size Standards*, *supra* note 148.

This Comment offers an alternative. By combining a proportional liability framework with sandbox innovation, it shows how accountability and experimentation can work together. Startups can test new systems in supervised environments, learning and adapting without the crushing weight of full compliance, while larger companies with more resources face stricter obligations to monitor, disclose, and remediate harms. This model keeps the playing field open and ensures the public is protected and gives regulators the chance to learn alongside industry rather than only reacting after the fact.

However, to make that vision workable, piecemeal tweaks to Section 230 are not enough. The law was never designed for systems that generate content on their own and trying to retrofit it risks either leaving critical gaps or smothering new entrants. What is needed instead is a modern federal framework: one that recognizes the differences between generative and predictive AI, scales obligations by company capacity and risk level, and embeds regulatory sandboxes as a mechanism for safe experimentation under clear guardrails. This approach preserves the spirit of Section 230's pro-innovation stance while crafting rules that reflect the realities of today's AI ecosystem. The choices lawmakers make in the next few years will determine whether the U.S. remains at the forefront of AI innovation while safeguarding consumers or whether the field consolidates under a handful of dominant players.