ALM | LAW.COM

New York Law Journal

Generative AI: The Next Frontier for Section 230 of the Communications Decency Act

By Avi Weitzman and Jackson Herndon
June 26, 2023

Generative Artificial Intelligence ("AI") is one of the most exciting innovations in our lifetimes. The technology has captured the imagination of millions around the world, and its adoption is spreading at lightning speed.

In January 2023, ChatGPT became the fastest growing consumer application of all time, reaching 100 million monthly active users just two months after its launch. (Krystal Hu, *ChatGPT sets record for fastest-growing user base*, Reuters, February 2, 2023, available at https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/). OpenAI, the company that released ChatGPT, proclaims that in the next ten years, AI systems could "exceed expert skill level in most domains, and carry out as much productive activity as one of today's largest corporations." (*Governance of Superintelligence*, available at https://openai.com/blog/governance-of-superintelligence).

Goldman Sachs, for its part, has predicted that generative AI could raise global GDP by 7%—or $7 trillion—in the next ten years. (Goldman Sachs, *Generative AI Could Raise Global GDP by 7%*, April 5, 2023, available at https://www.goldmansachs.com/intelligence/pages/generative-ai-could-raise-global-gdp-by-7-percent.html)

Yet even generative AI's biggest fans are quick to acknowledge its risks, with many calling for a moratorium on development and more regulation. (*See, e.g.*, *Center for AI Safety's Statement on AI Risk*, available at https://www.safe.ai/statement-on-ai-risk; Ryan Tracy, *ChatGPT's Sam Altman Warns Congress That AI 'Can Go Quite Wrong,'* Wall Street Journal, May 16, 2023, available at https://www.wsj.com/articles/chatgpts-sam-altman-faces-senate-panel-examining-artificial-intelligence-4bb6942a).

Credit: GarryKillian/Shutterstock

The risks are big and small. While some long term risks are potentially existential, there are also substantial, albeit less consequential, litigation risks to companies involved in generative AI. Those legal risks are only now starting to come into focus, and courts will grapple with them for years because, in part, even the creators of generative AI don't fully

understand how the AI programs generate their content.

As companies continue to jump into the generative AI bandwagon, they should recognize the risk that Section 230 of the Communications Decency Act ("CDA")—the principal statute that has served to protect and foster the internet over the past 25 years—may be deemed not to protect certain generative AI.

Section 230, adopted in 1996, immunizes web companies from liability for hosting third party content on their platforms; it has permitted those platforms to grow into the massive online marketplaces they are today. While the potential malicious use of generative AI to spread misinformation, disinformation, and defamation has been a hot topic of discussion, the legal press has paid little attention to the gatekeeping question of whether generative AI is protected under Section 230 of the CDA. This article explores that very issue.

What is Generative AI?

Generative AI is a subfield of AI in which computer algorithms are used to generate outputs that resemble human-created content, such as images, videos, art, music, text, and software code. The output is based on training data from large models—often containing millions of images, sentences, and/or sounds—from which a computer can learn to create the desired output.

Today, the most widely used AI tools are chatbots based on language models, like ChatGPT, whereby software can mimic human conversation with the user, as well as offer a range of functionalities, from writing computer code to composing stories, poetry, and song lyrics.

Generative AI also has visual analogs. These include text-to-image generators like Stable Diffusion, DALL-E, and Midjourney, each of which can quickly produce AI-generated images based on text prompts. For example, a user can ask Midjourney to generate a painting of lawyers fighting in court in the style of Vincent Van Gogh, and it will provide mock ups within seconds. Generative AI is now also quickly spreading into audio and video generation.

All of this, however, comes with potentially significant legal risk. One such risk involves defamation. For instance, on June 5, a pro-gun radio host in Georgia filed a defamation claim against OpenAI alleging that ChatGPT falsely accused him of embezzlement when a reporter asked ChatGPT to summarize the allegations in a pending Second Amendment lawsuit. (Isaiah Poritz, *First ChatGPT Defamation Lawsuit to Test AI's Legal Liability*, Bloomberg Law, June 12, 2023, available at https://news.bloomberglaw.com/ip-law/first-chatgpt-defamation-lawsuit-to-test-ais-legal-liability).

The allegations in that case involve the well-known phenomenon of ChatGPT hallucinations—where the software authoritatively states a false fact as truth. Most recently, such hallucinations captured legal headlines when a lawyer submitted a brief containing made up cases and quotes generated by ChatGPT in response to the lawyer's queries for legal authority. (*See* Benjamin Weiser, *Here's What Happens When Your Lawyer Uses ChatGPT*, New York Times, May 27, 2023, available at https://www.nytimes.com/2023/05/27/nyregion/avianca-airline-lawsuit-chatgpt.html?smid=nytcore-ios-share&referringSource=articleShare).

But the risks do not stop there. In fact, one of the most serious concerns is what bad actors might do with these new tools, including quickly spreading dis- or mis-information, as well as photographic, audio, and video deep fakes online.

For example, following Donald Trump's indictment in New York in March, realistic images of Mr. Trump being arrested, tackled, and carried away by police filled Twitter. (*See* Ashley Wong, *Paparazzi Photos Were the Scourge of Celebrities. Now, It's AI*, Wall Street Journal, April 3, 2023, available at https://www.wsj.com/articles/ai-photos-pope-francis-celebrities-dfb61f1d). None

of these images was real; they were created in seconds with image generative AI software.

Potential Section 230 Issues for Generative AI

Section 230 of the CDA was enacted in 1996, during the early days of the internet. Congress passed the statue with the express intent "to promote the continued development of the Internet and other interactive computer services." 47 U.S.C. §230(b)(1).

Section 230 immunizes (1) providers and users of "interactive computer service[s]" from being held liable (2) "as the publisher or speaker" (3) "of any information provided by another information content provider." 47 U.S.C. §230(c)(1); *see also* 47 U.S.C. §230(f)(3) ("The term 'information content provider' means 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.").

While the law has some exceptions—most notably for violations of criminal and intellectual property law—it broadly prohibits civil claims where a website republishes third party content. 47 U.S.C. §230(e)(1)-(3).

But the Internet has greatly evolved over the past 25 years, and in recent years, Section 230 has come under attack. Among other things, detractors contend that the law has protected undesirable, illegal, and dangerous online platforms, with lawsuits accusing web companies of everything from promoting sex trafficking to terrorism. *E.g.*, *Jane Doe No. 1 v. Backpage.com, LLC*, 817 F.3d 12, 29 (1st Cir. 2016); *Force v. Facebook*, 934 F.3d 53, 57 (2d Cir. 2019).

The U.S. Supreme Court has never issued a ruling defining the scope of Section 230. Indeed, although it granted *certiorari* in two Section 230 cases this past year—in *Gonzalez v. Google*, 598 U.S. __ (2023), and *Twitter v. Taamneh*, 598 U.S. __ (2023), where the plaintiffs sought to hold technology companies liable for aiding and abetting terrorism—the Supreme Court ultimately punted on the Section 230 questions.

Instead, writing for the majority, Justice Clarence Thomas held that the plaintiffs failed to state viable aiding and abetting claims against the tech companies because there was a "lack of any concrete nexus between" the defendant companies and the terrorist attacks. The majority opinion noted that the companies' "relationship with ISIS and its supporters appears to have been the same as their relationship with their billion-plus other users: arm's length, passive, and largely indifferent."

Justice Thomas also noted that the "mere creation of" social-media platforms that "bad actors" used for "illegal—and sometimes terrible—ends" does not impose liability on the tech companies, because "the same could be said of cell phones, email, or the internet generally." Based on the allegations in these lawsuits, the tech companies thus could not be held liable as aiders and abettors of terrorist attacks.

That more-passive view of the internet may change with the advent of generative AI. Given its ability to generate content, as its name suggests, there will be substantial litigation as to whether AI companies can seek Section 230 immunity.

The First Two Elements Under Section 230

To obtain Section 230 immunity, the first two elements a defendant must establish are that it is a "provider or user of an interactive computer service" and that the plaintiff is seeking to hold the defendant liable as a "publisher or speaker." We assume both of these elements will easily be met in many cases.

First, we do not expect that there will be a serious challenge for internet-based AI companies to establish themselves as providers of interactive computer services or for anyone who uses an AI platform to be deemed a "user" of an interactive computer service.

Second, courts have broadly defined what it means to be a publisher of information provided by a third party. This includes suits that attack

the underlying program's use of algorithms rather than just the publication itself. *Facebook*, 934 F.3d at 66 (rejecting plaintiffs' arguments that Facebook's algorithms remove Facebook's status as a publisher under Section 230).

Thus, we'll assume in this article that a plaintiff is either alleging that the generative AI defendant is the speaker or publisher of the statement.

The Third Element: Is Generative AI a Content Creator?

The crux of the Section 230 analysis will likely focus on the third element: whether the defendant is a content creator or merely providing third party information. This analysis will necessarily be fact intensive and turn largely on what both the user and the generative AI platform are actually doing to produce content.

To begin, many people who have spent even a few minutes with generative AI tools find the technology genuinely creative and truly transformative. Ask ChatGPT to write a short, witty poem about the Denver Nuggets in Iambic pentameter, or prompt Midjourney to create a portrait of President Biden in the style of Pablo Picasso—the results are amazing and seemingly novel. It is hard to imagine that these results are republishing content that already exists. This view of generative AI might place it outside the protection of Section 230.

While a facially appealing argument, this does not conclude the analysis. Generative AI companies will contend that their technology is solely driven by third party content, namely: (i) the person who prompts the machine, and (ii) whatever third party data the model was trained on.

Under this telling, tools like ChatGPT are just remixing third party content based on, and in response to, prompts from the actual creator of the content—the user of the generative AI. If that is correct, it will be argued that the technology is no different than other Section 230-immunized platforms that merely provide snippets of third party information in search results. *See, e.g.*, *O'Kroley v. Fastcase, Inc.*, 831 F.3d 352, 355 (6th Cir. 2016) (holding that Section 230 protects Google's automatically generated snippets that summarize search results).

That said, something else and more significant does appear to be happening with some generative AIs, more akin to the act of creation. And that act of creation, in whole in or in part, may change the Section 230 analysis. The U.S. Court of Appeals for the Ninth Circuit held in *Fair Housing Council of San Fernando Valley v. Roommates.com*, 521 F.3d 1157 (9th Cir. 2008), that Section 230 immunity may not apply when the website is responsible at least in part for the created content and the content is unlawful.

In that case, the defendant helped create subscribers' profile pages through questions the website asked the user to complete, including discriminatory questions about sexual orientation, which resulted in the subscriber being induced to express illegal preferences. *Id.* at 1164-67. Thus, where a website "helps to develop unlawful content" and "contributes materially to the alleged illegality of the conduct," it loses Section 230 protection. *Id.* at 1168.

Notably, though, while the Court found that Roommates.com could be held liable for its discriminatory questions, it also held that the website could not be held responsible for content its subscribers posted in a blank "Additional Comments" field, as those comments came "entirely from subscribers and [are] passively displayed by" the website. *Id.* at 1173-74. One year after *Roommates*, the U.S. Court of Appeals for the Tenth Circuit similarly held in *FTC v. Accusearch Inc.*, 570 F.3d 1187, 1199 (10th Cir. 2009), that "a service provider is 'responsible' for the development of offensive content . . . if it in some way specifically encourages development of what is offensive about the content."

The *Roommates* and *Accusearch* rulings thus paint a different picture of websites generating content than the one the Supreme Court analyzed in the *Twitter* case. Today, many generative AI

platforms only offer their users a blank box, like the "Additional Comments" field in the *Roommates* case, and whatever transpires is driven primarily by what the user chooses to enter as an initial prompt.

The question, then, is how will courts react to a generative AI that ***obliges*** a user's requests to engage in unlawful conduct—whether it's the creation of a defamatory or tortious deep fake video or an instruction manual for carrying out a violent attack? Will that be viewed as the AI tool taking a passive role, like providing a communications platform that allegedly enabled terrorist attacks, or is it actively permitting, if not encouraging, the use of technology for unlawful means? The former views generative AI as no different than any other technological tool—cell phones, email, or the internet generally—that can been used for illegal conduct, just as Justice Thomas found in *Twitter*; the latter risks the opposite.

None of this, however, fully answers the question of how to handle hallucinatory AI. Recall, again, the attorney who relied on ChatGPT to do his legal research only to discover that he had cited a number of completely fake authorities in a federal court filing. While it may be tempting to hold AI products responsible for inaccurate information, Section 230 has protected internet companies for displaying false and fraudulent information for years.

For example, eBay secured Section 230 immunity over 20 years ago in *Gentry v. eBay*, 99 Cal. App. 4th 816, 831 (Cal. Ct. App. 2002), when it was sued after some of its users posted and sold fake sports memorabilia on its platform. Particularly with proper disclaimers by the generative AI companies alerting users to confirm the information provided, it may be argued that users should bear the responsibility to ensure the accuracy of information generated by AI—such as the lawyer who submitted the fake authorities to the court. In the meantime, whatever the Section 230 result may be, generative AI companies should continue to invest in figuring out solutions to the hallucinatory aspects of generative AI.

Finally, we note that the staggering speed of development may be challenging for courts analyzing this issue. Already, some generative AI does not even need direct user prompts to generate content. Open source projects like Auto-GPT, for instance, now prompt themselves to complete multi-step tasks. (Kyle Wiggers, *What is Auto-GPT and Why Does It Matter?*, TechCrunch, April 22, 2023, available at https://techcrunch.com/2023/04/22/what-is-auto-gpt-and-why-does-it-matter/). In this situation, the chatbot may not be providing information based on third party prompts but, instead, arguably based on its own prompts.

The courts will no doubt find it difficult to keep up with these technological advances, and Section 230 itself may prove an inadequate tool for the regulation of this brave new frontier of generative AI. The weighty questions—and risks—now presented by generative AI may very well be the impetus for a new or revised statutory regime.

Indeed, two U.S. senators have already proposed one bill on the matter that would waive Section 230 immunity for all generative AI. (//www.hawley.senate.gov/sites/default/files/2023-06/Hawley-No-Section-230-Immunity-for-AI-Act.pdf).

Given the wide variety of opinions regarding the application of Section 230 to the internet more generally, we expect that Congress will be debating this issue with respect to generative AI for quite some time to come.

Avi Weitzman, *a former Assistant U.S. Attorney in the Southern District of New York, is a partner at Paul Hastings, where he is co-chair of the firm's complex litigation and arbitration practice.* **Jackson Herndon** *is a senior associate at the firm.*