

Section 230 Immunity and Generative Artificial Intelligence

December 28, 2023

Over the past year, technology companies have expanded access to services capable of creating content using artificial intelligence (AI). In February 2023, Microsoft [announced](#) an “all new, AI-powered Bing search engine” that “can help you write an email, create a 5-day itinerary for a dream vacation to Hawaii, with links to book your travel and accommodations, prep for a job interview or create a quiz for trivia night.” Google introduced [Bard](#), an “experimental conversational AI service,” the same month. Applications such as these that are capable of generating new content like text, images, and videos form a subset of AI applications often referred to as “generative AI.”

As access to AI tools has expanded and the underlying models have become more powerful, Congress has shown interest in regulating AI models. Committees [have](#) held [hearings](#), and Members have [introduced bills](#) to regulate AI (or announced a [framework](#) for doing so). Some Members have [called](#) for a [task force](#) or [commission](#) to recommend rules. The executive branch has also weighed in. The White House issued an [Executive Order](#), in October 2023, tasking agencies with actions to [address](#) the advent of AI models.

Generative AI poses unique policy and legal issues. One question raised by the introduction of generative AI products is the extent to which companies that provide the products could be held liable for illegal content generated by the AI. The answer likely depends in part on an existing legal framework: [Section 230 of the Communications Act of 1934](#), a federal statute that, subject to some exceptions, immunizes interactive computer service providers from being sued as the publisher or speaker of information provided by another party. If this immunity extends to claims based on an output from a generative AI product, plaintiffs defamed by an AI output (for example) may be barred from suing a company that provided the AI product. The potential application of Section 230 to generative AI has already garnered comments at a Supreme Court [oral argument](#) and from Section 230’s primary [authors](#). This Sidebar discusses how Section 230 might apply to legal claims involving generative AI products.

Section 230

Section 230 creates a federal immunity for publishing another person’s content online, as explained in a [longer CRS report](#). Specifically, [Section 230](#), enacted in the Communications Decency Act of 1996, prevents providers and users of “interactive computer services” from being held liable—that is, legally responsible—as the “publisher or speaker” of information provided by another person. “Interactive

Congressional Research Service

<https://crsreports.congress.gov>

LSB11097

computer service” is [defined](#) as any service that “provides or enables computer access by multiple users to a computer server.” This broad term encompasses services such as [Facebook](#), [Google](#), and [Amazon](#). There are [exceptions](#) to Section 230 immunity, such as for intellectual property law.

Section 230 only provides immunity for content provided by *another* person and does not apply if a provider or user of an interactive computer service helped create or develop the content. Although the Supreme Court has never interpreted Section 230, [many federal and state courts](#) have considered when a lawsuit would attempt to hold a defendant liable for another’s content or for the defendant’s own content. A number of courts have settled on a [“material contribution” test](#), under which Section 230 immunity does not apply if the provider or user materially contributed to the alleged unlawfulness of the content. This inquiry is highly fact-specific. The U.S. Court of Appeals for the Ninth Circuit, for example, decided in [one case](#) that Section 230 barred housing discrimination claims against the operator of the website Roommates.com based on one portion of the website but did not bar claims based on a different portion of the site. The court [said](#) an open text field where roommate seekers could describe “what [they] are looking for in a roommate” did not materially contribute to discrimination, even if users’ descriptions ended up facilitating discriminatory searches. In contrast, the court [held](#) that “search and email systems [designed] to limit the listings available to subscribers based on sex, sexual orientation and presence of children” did materially contribute to alleged discrimination. The court [suggested](#) sites could provide “neutral tools” to post user content so long as the tools do not materially contribute to illegal activity.

Citing this “neutral tools” analysis, some federal courts of appeals have held that Section 230 [barred lawsuits](#) against service providers for promoting harmful content through algorithms, where the algorithms used objective factors that treated the harmful content similarly to other content. Beyond [content recommendation algorithms](#), courts have [held](#) that other choices about how to present material—and even actions that rearrange or slightly [edit material](#)—can also be immunized by Section 230. In *O’Kroy v. Fastcase, Inc.*, a federal appeals court concluded that Section 230 barred a defamation lawsuit challenging the way Google presented its search results. The plaintiff alleged that when he searched his name, one entry on the search results page inaccurately suggested that he had been involved in a case regarding indecency with a child. The court [held](#) that even though Google “performed some automated editorial acts on” content provided by other parties, “such as removing spaces and altering font,” these alterations did not materially contribute to the alleged defamatory nature of the content.

In contrast, in *FTC v. Accusearch Inc.*, a federal appeals court held Section 230 immunity did not apply because the website operator that was sued had helped to develop third-party information. The website sold information contained in telephone records. Although the website acquired the records from other parties, it solicited and paid for them. The legal harm came from the [publication](#) of the records: the FTC alleged the website operator [violated](#) a federal provision limiting disclosure of certain personal information. The court [said](#) the website operator was responsible for developing the information and thus not protected by Section 230, as it “contributed mightily to the unlawful conduct” by paying researchers to acquire confidential records with the intent of making them public contrary to federal law.

Even if Section 230 does not apply, a service provider will only be liable in a lawsuit if a plaintiff can prove their underlying legal claims. For example, if a copyright lawsuit falls within the Section 230 exception for intellectual property law, the plaintiff will still have to prove the provider in fact violated copyright law. Other CRS products may be relevant to that inquiry. For example, other products provide [an introduction to tort law](#) and discuss generative AI and [copyright law](#) or [campaign advertising](#).

Generative AI

“Artificial intelligence” is a broad term referring to computerized systems that act in ways commonly thought to imitate human intelligence. Generative AI is a type of AI that uses machine learning to generate new content, as discussed in [this In Focus](#). For instance, a user can provide a text prompt that the

generative AI uses to create an image or other content. [Machine learning](#) algorithms allow computers to learn at least partially without explicit instructions. Programmers gather and prepare training data, then they feed that data into a machine learning model. There are [different types of generative AI](#), but applications based on large language models, such as OpenAI's ChatGPT, were trained by providing the model with massive amounts of existing data scraped from the internet. ([This CRS report](#) discusses some of the data privacy concerns with generative AI.)

Accordingly, when generative AI creates new content, it [draws](#) from this existing information created by other parties and attempts to match the style and content of the underlying data. However, the new content may not be identical to the training data. Generative AI may sometimes produce so-called “[hallucinated](#)” outputs that, for instance, go beyond the scope of the training data or have incorrectly decoded that data. The generative outputs are also [influenced](#) by the specific text of the prompt itself—because the algorithms generate content based on statistical probabilities that the underlying training data are associated with the words in the prompt, changes to the words in the prompt can change the output. Even the same prompt provided to a generative AI system multiple times can result in different outputs.

Potential Application of Section 230 to Generative AI

Courts have not yet decided whether or how Section 230 may be used as a defense against claims based on outputs from recently released generative AI products, but they may soon be asked to address the issue. Outputs from generative AI products have already led to lawsuits. In June 2023, for example, a radio host [sued](#) OpenAI, alleging that ChatGPT defamed him. In a [second](#) defamation case, a plaintiff alleged that searching for his name on Microsoft's Bing search engine returned an AI-generated summary that commingled facts about him with facts about a different individual who had a similar name and who once pleaded guilty to seditious conspiracy. Although no defendant in these cases has yet raised Section 230 as a defense, the claims are similar to those that have given rise to Section 230 defenses in the past. In each case, the plaintiff alleges that a defendant published information using a tool that relies at least in part on user inputs and [data created by other parties](#). At least one commentator has [asserted](#) that generative AI companies might attempt to invoke Section 230 in similar circumstances. Even if raised, the courts deciding these cases may not address the issue. The cases could settle or be resolved on the merits or another defense. Still, these or future lawsuits alleging that generative AI outputs are defamatory, negligent, or cause other legal harms may test whether the providers of the tools or users who requested those outputs materially contributed to the harmfulness of the challenged content such that Section 230 immunity would not apply.

If Section 230 were to be invoked in such circumstances, past cases suggest Section 230 would be applied in a fact-specific manner. Thus, a Section 230 inquiry in a lawsuit challenging any given AI-generated output would likely depend on the particular legal claim and underlying facts. Not all generative AI products function the same way, and not all legal claims would necessarily turn on the same aspects of a given product. As one [group of scholars asserts](#), generative AI products “operate on something like a spectrum between a retrieval search engine (more likely to be covered by Section 230) and a creative engine (less likely to be covered).” Section 230's application could therefore vary across different generative AI products, different applications of a single product, and different legal claims about an application or product.

As such, commentators have made different predictions about how courts would resolve Section 230 defenses based in part on which aspects of generative AI the arguments focus on. Some commentators [contend](#) that “AI programs’ output is composed by the programs themselves,” so the AI providers should be viewed as information creators or developers that receive no Section 230 immunity. Large language models, for example, [can](#) “draft[] text on a topic in response to a user request or develop[] text to summarize the results of a search inquiry.” A [commentator](#) has argued that this aspect of the technology would likely lead courts to conclude that large language models develop text themselves. A large

language model's output can "hallucinate," creating brand new "text on a topic" that no other party has ever written. If generated content contains claims or assertions that do not appear in its training data, the claims or assertions could be seen as entirely new information created by the providers rather than by another person. Even if an AI program assembles facts or claims from training data into new material that does not appear elsewhere on the internet, this may be viewed as similar to *Accusearch*, where the website operator was deemed "responsible for the development of the specific content that was the source of the alleged liability" and thus was unprotected by Section 230.

On the other hand, another commentator *argues* that the current iteration of ChatGPT (as of March 2023) "is entirely driven by third-party input" and "does not invent, create, or develop outputs absent any prompting from an information content provider." When it receives a prompt, ChatGPT "*uses* predictive algorithms and an array of data made up entirely of publicly available information online to respond to [the] user-created inputs." Outside the context of Section 230, OpenAI has *claimed* that the "content" associated with any particular ChatGPT interaction includes both the machine-generated outputs and the user-generated inputs that prompted them. Courts focusing on these aspects of the product could consider ChatGPT analogous to search engines that use algorithms to generate lists and illustrative snippets of websites in response to user inputs. As discussed, federal appeals courts have held that Section 230 shields internet search engines from liability for claims based on the *results* returned by users' *searches*.

Autocomplete features are another potential analogy, given that some generative AI *operates* by predicting and assembling plausible text outputs that it associates with user prompts and patterns learned from training datasets. Although there is relatively little case law on point, *two* federal trial *courts* have held that Section 230 immunizes search engines from claims that auto-generated, suggested search terms were defamatory. The courts *reasoned* that the "auto-generated terms 'indicate[] only that other websites and users have connected plaintiff's name' with certain terms," meaning that the allegedly harmful content was created by another party. This reasoning could *apply* to an AI product that completes a sentence by "predict[ing] probabilities for the next plausible word or phrase given previous words or phrases."

Similarly, some courts have *held*, under the *neutral tools test*, that other algorithmic processes do not develop content for the purposes of Section 230. For example, the U.S. Court of Appeals for the Second Circuit held that Facebook does not develop or create content that its algorithms arrange and display on a user's page when the algorithms "*take* the information provided by . . . users and 'match' it to other users . . . based on objective factors applicable to any content." As mentioned, a large language model *can be* "trained simply to predict probabilities for the next plausible word or phrase given previous words or phrases." Some scholars have *argued* that this sort of probabilistic compilation *could* likewise be considered a tool that operates neutrally, regardless of the particular content to which it is being applied. The same algorithmic process that generates an output containing defamation or other illegal content in response to some user interactions could generate lawful outputs in response to other interactions. The neutral tools test, however, has been criticized by *some*, including with respect to how it might apply to AI products. A court could deny Section 230 immunity without resorting to the neutral tools test if it concluded in a given case that an AI output was created by the AI provider rather than information created by *another* person.

In all of these potential analyses, details will matter. Regardless of how the contours of the analysis are ultimately defined, current case law suggests that courts applying Section 230 to claims involving generative AI products would need to look closely at how the particular AI product at issue generates an output and what aspect of the output the plaintiff alleges to be illegal. Because generative AI products are not all the same, are likely to continue to evolve, and can rely on data and inputs from disparate sources, Section 230 analysis may lead to different outcomes in different cases.

Considerations for Congress

As Congress contemplates regulating generative AI, it may consider whether any proposals that would impose liability on generative AI companies would potentially conflict with Section 230 immunity. If so, Congress might consider creating Section 230 carveouts in any new legislation. Without such a carveout, courts may construe a law that imposes civil liability on generative AI companies to be limited by Section 230. For instance, Congress could create a new private right of action for harms stemming from generative AI outputs. The new right of action could define the elements a plaintiff would have to prove to establish liability and could specify whether Section 230 provides a defense. Congress might also consider creating new immunities to legal claims that relate to generative AI outputs. A new immunity could preclude liability regardless of whether Section 230 may also apply.

In addition, Congress could amend Section 230 to address generative AI directly. This sort of amendment could extend immunity to claims arising from generative AI products, withhold immunity from such claims, or extend immunity to certain types of claims and withhold it from others. A bill introduced in the 118th Congress (S. 1993), for example, would withhold Section 230 immunity “if the conduct underlying the claim or charge involves the use or provision of generative artificial intelligence by the interactive computer service.” This bill would remove immunity not only for generative AI providers but also in any claim *involving* a provider’s *use* of generative AI. Alternatively, Congress could wait to see how courts address defenses under the current version of Section 230 in cases based on outputs from generative AI.

Other proposals to more broadly amend Section 230 could also have implications for generative AI. For instance, the DISCOURSE Act (S. 921) would effectively create an exception for certain service providers that modify or alter others’ information. This approach could sweep in generative AI that edits third-party content in creating new outputs.

Congress’s ability to regulate in this area could be limited by the [First Amendment](#)’s Free Speech Clause. Some scholars have [argued](#) that generative AI output is protected by the First Amendment, either because the creators use the program as a means of expression or because the program’s users have a right to use generative AI as a method of creating or receiving expression. The Supreme Court has [ruled](#) that “the creation and dissemination of information are speech within the meaning of the First Amendment,” such that even factual information (and sometimes [false information](#)) can be constitutionally protected. These cases suggest a law creating or allowing liability for using generative AI could implicate free speech interests. If Section 230 did not bar a lawsuit, whether under existing interpretations of the law or because Congress created a new exception, the protections of the First Amendment [might still apply](#).

A law that affects speech is not necessarily unconstitutional, however. First Amendment protections are not absolute. Instead, courts apply different types of constitutional scrutiny depending on the type of speech being regulated and how the regulation affects that speech. A law amending when Section 230 immunity applies might be [scrutinized differently](#) than a law more directly regulating generative AI activity. Further, for example, laws that regulate speech [based on its content](#) are generally subject to stricter constitutional scrutiny than [content-neutral regulations](#). Thus, the constitutionality of a law regulating generative AI, including through an amendment to Section 230, may depend in part on [whether](#) the law targets the function of a generative AI program or its expressive output and on whether it targets specific types of AI outputs based on their content.

Author Information

Peter J. Benson
Legislative Attorney

Valerie C. Brannon
Legislative Attorney

Disclaimer

This document was prepared by the Congressional Research Service (CRS). CRS serves as nonpartisan shared staff to congressional committees and Members of Congress. It operates solely at the behest of and under the direction of Congress. Information in a CRS Report should not be relied upon for purposes other than public understanding of information that has been provided by CRS to Members of Congress in connection with CRS's institutional role. CRS Reports, as a work of the United States Government, are not subject to copyright protection in the United States. Any CRS Report may be reproduced and distributed in its entirety without permission from CRS. However, as a CRS Report may include copyrighted images or material from a third party, you may need to obtain the permission of the copyright holder if you wish to copy or otherwise use copyrighted material.