April 7, 2026

**EXPERT INSIGHT** | AI and Emerging Technology

# AI Chatbot Regulation and the First Amendment

*Joel Thayer*

## TOPLINE POINTS

- Tech companies argue the First Amendment shields AI chatbots from regulation, but recent Supreme Court precedent — including *Paxton, Moody,* and *TikTok v. Garland* — significantly weakens that claim, especially for child safety measures.
- Unlike social media, chatbots are not "modern public squares." They function more like interactive services that respond to individual users, giving them far weaker First Amendment protections and opening the door to commonsense regulations like account creation, age verification, parental oversight, and transparency requirements.
- Legislatures have a strong constitutional basis to require user accounts, age verification, custodial controls, and transparency disclosures for generative AI platforms without running afoul of the First Amendment.

## INTRODUCTION

The America First position on AI is simple: ensure that America leads the world in developing and adopting AI for human flourishing. An essential part of that agenda is protecting children. This has led many state and federal legislators to decide on a wide array of solutions to address the concerns many of these commercial-grade apps raise, particularly the impact they have on children. From light-touch approaches, like age verification, design codes, and transparency requirements, to banning aspects of the technology, legislatures are introducing and enacting these measures.

As in all cases concerning legislative prescriptions involving tech, tech companies and their proxies have raised standard First Amendment defenses to thwart any regulation over these systems. For example, one argument suggests that AI applications (e.g., large language models, or "LLMs") are akin to a library or a bookstore, and that, therefore, the government cannot forbid the use of these applications. But are these apt analogies?

It is true that users can get a great deal of information from these applications. It is also true, however, that these applications can pose severe or even fatal consequences. Unlike a bookstore or library, a chatbot can contextualize information to encourage a child to fall in love with it, while asking them to conceal information from their parents. Much worse, as we are continuing to find out, they have even assisted kids with committing suicide and with their suicide attempts. These are hardly the types of interactions a child will encounter with a librarian or at a Barnes and Noble bookstore.

What sort of speech would we be protecting? And who is speaking? Importantly, does the First Amendment even protect these types of communications? The goal of this insight is to address these questions head-on.

## The state-of-play for the first amendment

The Free Speech Clause of the First Amendment states that "Congress shall make no law . . . abridging the freedom of speech[.]" The scope of this right is hotly debated, and the case law is, in a word, "messy." At the founding, "the Freedom of Speech" referred to a natural right held by individuals to be free from prior restraints, laws that restricted publications, and criminal punishment imposed by the State, save those based on seditious libel. However, more modern cases have augmented that view to protect a whole host of activities from particular government restraints. For instance, it prohibits the government from requiring newspapers to host views contrary to its own. In the tech context, the Supreme Court has even forbidden states from broadly restricting adults' access to social media platforms.

But recent Supreme Court cases indicate that jurisprudence is moving from the more quasi-libertarian perspective—that defines the First Amendment's scope broadly—to a narrower view steeped in originalism—the idea that modern interpretations must be rooted in concepts conceivable by the founding generation. For example, *Free Speech Coalition v. Paxton* upheld a Texas law that applied age verification to porn websites. Although many commenters argue that the decision is narrowly focused on content obscene to minors, the decision's expressed limits are arguably much broader.

Citing the Federalist Papers, Justice Clarence Thomas writes that "'[n]o axiom is more clearly established in law, or in reason, than that . . . wherever a general power to do a thing is given, every particular power necessary for doing it is included.'" According to the Supreme Court, "where the Constitution reserves a power to the States, it also reserves "the ordinary and appropriate means" of exercising that power." In other words, if the government is permitted by law to place restrictions on a product and service in the real world, then it may do so on its digital counterpart. That involves much more than regulating obscenity only.

Another example is *Moody v. NetChoice*, where the Court considered whether two laws from Florida and Texas that regulated social media platforms were unconstitutional. The Supreme Court concluded that the laws had too many different working parts and applications to make a sweeping decision one way or another. They remanded the cases to the lower courts, chastising the plaintiffs for using a broad facial challenge instead of particularized, as-applied challenges. In *Moody*, the Court generally

punted the First Amendment question, but offered some helpful *dicta* (i.e., non-binding language) throughout the opinion. In *Moody*, Justice Kagan (citing Justice Barrett) stated that the case was limited and does not apply to "feeds whose algorithms respond solely to how users act online[.]" That matters for AI systems, which mainly generate responses to user prompts. While the Court did not create a clear rule, its reasoning suggests that AI platforms may have weaker First Amendment protection when they function more like tools than speakers.

We are starting to see lower courts apply Justice Barrett's logic to AI with *Garcia v. Character Technologies*. The case involves the death of 14-year-old Sewell Setzer III. Sewell's mother, Megan Garcia, argues that Sewell's use of the app Character AI caused his death. The facts of the case are compelling. Sewell downloaded Character AI on his mobile phone on April 14, 2023. On the app, he interacted with chatbots posing as "licensed CBT therapists" and a variety of Game of Thrones characters. Within a couple of months, Sewell became addicted to the app, which led to him spending more time secluded in his room and contributed to him quitting the junior varsity basketball team. Things got worse after he upgraded his account to Character AI's paid subscription that provides an even more enhanced experience for users. After that purchase, Sewell's mental state and school performance continued to rapidly decline. On his therapist's recommendation, Sewell's parents confiscated his phone "until the end of the school year," but to no avail. Sewell found it and sent his last message to the Daenerys Targaryen character on the app. Soon after that message, he placed a gun to his head and took his own life on February 25, 2025.

On behalf of her son's estate, Megan Garcia sued Character AI under two tort theories (i.e., wrongful death and negligence) and claimed that the app engaged in unfair and deceptive practices. Character AI motioned to dismiss the case, *inter alia*, because the court granting Garcia's relief would offend its users' First Amendment rights. Judge Anne Conway denied Character AI's motion to dismiss, especially on those grounds. Judge Conway, quoting Justice Barrett, explained that "a platform creates an algorithm to remove posts supporting a particular position from its social media site, "the algorithm [] simply implement[s] [the entity's] inherently expressive choice 'to exclude a message.'"" Judge Conway contrasted this, stating that "[t]he same might not be true of A.I. though—especially where the A.I. relies on an LLM[.]" She questioned, as did Justice Barrett, whether there is any expression when a platform "hands the reins" over to an AI to make curation decisions. Thus, she was "not prepared to hold that Character A.I.'s output is speech" and allowed the case to continue to trial.

The last major case is *TikTok v. Garland*. In that case, TikTok challenged a law that bans ByteDance, its Chinese parent company, from owning TikTok in the United States, given the Chinese government's control over ByteDance and the national security and sovereignty risks that control creates. TikTok argued that the required divestiture imposed a "disproportionate burden upon" their First Amendment activities," which included its "content moderation" and "content generation." In that case, the Court appeared skeptical of TikTok's argument that a mere regulation of an algorithm raises First Amendment scrutiny. What is more, the Court held that "Even assuming [one of the law's] rationale[s] turn[] on content, … [t]he record before [them] adequately supports the conclusion that Congress would have passed the challenged provisions based on the data collection justification alone." Here, the Court clarified that a law regulating a tech platform does not necessarily invite a First

Amendment review if the law's primary justification is not content-based, even if ancillary justifications are.

The first thing to identify with content-generating AI (hereafter, "AI systems") is the precise speech interest in question.

## WHAT CAN BE DONE TO PROTECT KIDS WITHOUT OFFENDING THE FIRST AMENDMENT?

Given that most AI applications do not predominantly distribute obscenity, the most prudent route for any legislator (federal or state) is to steer away from content-based restrictions. Even still, that leaves a lot on the table, especially when imposing various child safety restrictions on AI systems. Given the limited rights to which AI systems are likely to enjoy, one of the most straightforward regulations would be one that requires users to create accounts before use. Safeguards like age verification, transparency requirements, and parental oversight would also be ripe for enactment under such a regulatory scheme.

## REQUIRE USER ACCOUNTS BEFORE USING A CHATBOT

Given that the tech companies know kids are accessing their services and the dangers towards kids are well established, all AI systems should require users to create accounts before they may use the service. Under *Paxton*, the government certainly has the power under traditional common law to impose obligations on property owners who know or should know that children will enter their property. Such a law need not be complicated. It would rest on over a century of established "attractive nuisance" precedent.

The attractive nuisance doctrine applies liability "when the owner or possessor of land maintains an artificial condition on her property that entices a child to enter her property and causes physical harm or injury." Typical "artificial conditions" in this context range from everything from a pool to a trampoline. In such a scenario, the law imposes a higher duty on property owners, given the relative danger that such a condition is known to have on a child. Such requirements include placing fences around the pool or locking up the trampoline when one is not at home.

Recent research finds that 64% of children are using chatbots—30% of those surveyed say they use them daily. This widespread use, along with the instances of child suicides as a direct result of using these services, ought to provide a rationale for requiring users to create accounts before engaging with a chatbot. It is not only substantially related to the important government interest of protecting children, but it presents almost no burden to adults' expressive activity (to the extent any exists).

Of course, there may be some who argue that this measure may require the user to relinquish personal data and, thus, violates their First Amendment right to access information anonymously. They may argue that such a requirement may cause users to not use the product for a variety of reasons.

However, the implied right to anonymity is usually applied to human-to-human interactions, and courts ascribe the "chilling effect" here primarily to a fear of "economic or official

retaliation" or "social ostracism[.]" Although these factors may exist on social media, where users engage in vigorous debate and share opinions, that is certainly not present when your only interaction is with a robot. Thus, the fear of being judged or ostracized by your peers or the public is substantially diminished or even nonexistent.

It is important to note that this area of law is still developing, and there are several appeals in process that may change the analysis. However, at this point, a worthy path for legislatures to pursue would be to have AI systems require their users to create an account before using them due to (a) the high risk they pose to children, (b) the government's interest in protecting children, and (c) a strong showing that this is the "most effective way of achieving its interest."

## REQUIRE AGE VERIFICATION AT THE TIME AN ACCOUNT IS CREATED

With an account-creation requirement at the front end, placing age verification requirements becomes far more manageable because, like any other establishment, they can check your ID at the door. As discussed above, the Court appears open to accepting restrictions on digital platforms that are within the purview of a state's power in a brick-and-mortar context. This includes applying age verification to certain services.

Under Court precedent, opponents of age verification for chatbots will have a tall order in establishing that a child has a First Amendment right to access the platform in the first place. To start, chatbots are unlike social media, and they are unlikely to raise the same speech interests. As the Court in *Packingham* explained, "users can debate religion and politics with their friends and neighbors," or "petition their elected representatives and otherwise engage with them in a direct manner" on social media. The same cannot be said about a chatbot, which serves more like an interactive search engine where responses are prompted by user queries. A chatbot is no more a "modern public square" than a child's diary.

But there is a slight complication. Based on the ruling in *Paxton*, lawmakers will first have to define what their traditional constitutional right to regulate is. One could be their right to prevent platforms from engaging in unfair or deceptive practices (UDAP) on child consumers. Given that children generally do not have the capacity to engage with commercial entities without parental oversight, the government could argue that the imposition of an age gate requirement on a platform is the "ordinary and appropriate means" of its constitutional power to protect child consumers from UDAP practices.

The Court's rationale in *Paxton* proves instructive. Indeed, Justice Thomas discusses how "[r]equiring age verification is common when a law draws lines based on age." He then points to examples of how Texas requires "proof of age to obtain alcohol, tobacco, a tattoo, [or] a body piercing[.]" What is more, he cites two federal examples of age-gated products and services, like the ability "to obtain certain medications from a pharmacist, … [or] employment as a minor." Justice Thomas went further still when explaining that states, specifically Texas, age-gated activities that implicate fundamental rights, such as issuing handgun licenses, registering to vote, or obtaining a marriage license. He closes this analysis by pointing out that "[i]n none of these contexts is the constitutionality of a reasonable, bona fide age-verification requirement disputed."

Even so, a court may find that such a restriction is tantamount (despite their obvious distinctions) to preventing users from accessing social media. If so, then a court may find that such a measure will "disproportionately burden[] expression" by "regulating who may speak and access speech on certain types of … platforms[.]" This means that the First Amendment would apply and that the regulation's level of scrutiny would depend on whether the government's regulation is triggered by the existence of specific content or limited to only some platforms.

If the regulation applies to all chatbots, then a court will likely hold that it is content-neutral and, thus, subject to intermediate scrutiny. In such an event, applying age verification has a high possibility of passing muster under the First Amendment, because the law feels that those types of regulations "pose a less substantial risk of excising certain ideas or viewpoints from the public dialogue." A court applying intermediate scrutiny "will sustain a content-neutral law 'if it advances important governmental interests unrelated to the suppression of free speech and does not burden substantially more speech than necessary to further those interests.'"

The government has an "important governmental interest" here, as protecting children is a well-established and accepted interest.

The question will rely on whether the imposition of disclosing age does not burden more speech than necessary. If so, the government will likely prevail. Indeed, the Supreme Court in *Reno v. ACLU* held that the primary issue with the government imposing its restrictions under the Communications Decency Act was that the platforms had "'no effective way to determine the identity or the age of a user who is accessing material through e-mail, mail exploders, newsgroups or chat rooms.'" However, at the time of that opinion, as Justice Sandra Day O'Connor opined, "screening software [was] 'not in wide use today' and 'only a handful of browsers [had] screening capabilities.'"

Justice O'Connor would be astonished at the extraordinary advances in today's age verification technology, especially with the assistance of AI. For instance, OpenAI has already started to create models to "predict" its users' ages who engage with its chatbot. Apple and Google are already obligated in other jurisdictions to age-verify the download of mobile apps in the U.K., which is primarily how children access chatbot applications. In fact, AFPI has offered app store age verification as a potential method to protect kids online. Thus, moving away from simple self-declaration toward more accurate, AI-driven, and privacy-preserving methods is not only possible, but no longer the burden it was when the Court decided *Reno*.

This is reflected in the caselaw as the Supreme Court now views age verification as a "modest burden" provided the government has an established constitutional right to act, especially with how pervasive today's technological advances are. As the Court explained, "[w]ith the rise of the smartphone and instant streaming, many adolescents can now access vast libraries of video content—both benign and obscene—at almost any time and place, with an ease that would have been unimaginable at the time of *Reno* and *Ashcroft II*." The rise of chatbots and their associated harms can only contribute to the Court's concern.

It is also important to note that, in *Packingham*, Justices Alito, joined by Chief Justice Roberts and Justice Thomas, the more originalist minority, disagreed with the majority's characterization of social media being akin to "a street or a park" for the purpose of the First Amendment. Given that the current justices on the Court lean towards a more originalist direction, it may be fair to assume the Court's view of social media as a "modern public square" has shifted; it also follows that the Court may not view chatbot platforms as public squares at all.

## REQUIRE CHILD ACCOUNTS BE ATTUNED TO A CUSTODIAN ACCOUNT

The implementation of custodial accounts (those managed on behalf of a minor) necessarily requires a child account. As such, requiring user accounts is a very useful tool for implementing custodial accounts. As to the First Amendment, if the Court retains its originalist posture, then it appears likely that the Supreme Court will take a refreshed look at laws requiring parental oversight, such as parental consent mechanisms or even requiring custodial accounts for chatbot use. Traditionally, a child's right to access information or right to expression has been tempered or enhanced by the adults' rights. As Justice Thomas outlined,

> "The historical evidence shows that the founding generation believed parents had absolute authority over their minor children and expected parents to use that authority to direct the proper development of their children. It would be absurd to suggest that such a society understood "the freedom of speech" to include a right to speak to minors (or a corresponding right of minors to access speech) without going through the minors' parents. … The founding generation would not have considered it an abridgment of "the freedom of speech" to support parental authority by restricting speech that bypasses minors' parents."

A child's First Amendment right is particularly limited when compared to the fundamental right to parent. As the Supreme Court acknowledged, "the child is not the mere creature of the AI State; those who nurture him and direct his destiny have the right, coupled with the high duty, to recognize and prepare him for additional obligations." Indeed, the Supreme Court has found that parents have a fundamental right "in the companionship, care, custody, and management" of their children. Further, the Court stated "[t]his primary role of the parents in the upbringing of their children is now established beyond debate as an enduring American tradition." Given this, the requirement that controls, such as filtering or time limits, be made available to parents may be perfectly acceptable.

But even if a court failed to implement a balancing test, it would likely rely on *Paxton*. In order to impose a custodial account requirement on AI systems' networks, the law must be an "ordinary and appropriate means" derived from the government's current constitutional power. In other words, Congress must regulate a digital service, in this case applications like ChatGPT and Grok, in the same manner the law allows it to for similarly situated traditional, non-digital services. However, as is the case with any new precedent, we are going to need more case law to develop to see precisely what the "ordinary and appropriate means" entails.

Even so, this framing suggests that the best paths forward would be to:

1. Keep the regulation content neutral given that chatbots' primary purpose, unlike the companies at issue in *Paxton*, is not to distribute content that is obscene to minors; and
2. Review brick-and-mortar products or services where we traditionally require custodial accounts or parental oversight.

Some services that either require custodial accounts or allow parental oversight that may serve as a basis for a custodial account requirement include:

- **Brokerage Accounts:** The law requires custodial accounts for holding stocks, bonds, and mutual funds, as minors cannot enter binding contracts for securities.
- **Banking/Cash Management:** The law requires custodial accounts for holding large cash gifts or inheritance intended for a minor that exceeds basic joint bank account restrictions.
- **Psychologists:** Parents generally have the right to consent to a child's mental health treatment and access records, but these rights vary by state and custody arrangements.
- **Educational Institutions:** Parents possess fundamental, constitutionally protected rights to direct the upbringing, education, and health decisions of their children, which extend to overseeing their interactions with teachers and school officials.
- **Libraries:** Parents have the right and responsibility to make decisions about what materials are suitable for their own family in a library.

## REQUIRE TRANSPARENCY

Irrespective of a requirement to create an account, legislatures could impose transparency requirements on AI systems. In general, courts evaluate First Amendment considerations concerning disclosure requirements under *Zauderer v. Office of Disciplinary Counsel*. The Supreme Court in *Zauderer* viewed disclosure requirements as content neutral because their purpose is to disclose "purely factual and uncontroversial information" about their conduct toward their users and the "terms under which [their] services will be available." The Supreme Court further clarified in *Milavetz, Gallop & Milavetz, P.A. v. United States* that while "restrictions on nonmisleading commercial speech regarding lawful activity must withstand intermediate scrutiny," it also held that for "provisions [that] impose a disclosure requirement rather than an affirmative limitation on speech . . . [rational basis] described in *Zauderer* governs [thei]r review." Additionally, a commercial disclosure requirement must be "reasonably related to the State's interest in preventing deception of consumers" and must not be "[u]njustified or unduly burdensome" such that it would "chill[] protected speech."

Even still, a court will likely consider a disclosure as a regulation of commercial speech. If that is the case, the lawmakers must: (a) identify a substantial governmental interest; and (b) demonstrate a sufficient fit between the law's requirement and that substantial government interest.

However, there are limitations. For example, lawmakers cannot use "deception" as a guise to compel the company to make certain political conclusions, which are inherently controversial and have no

objective standard on which to base the claims. Such disclosures may raise compelled speech concerns, warranting strict scrutiny review.

The Court has also found that some corporate disclosures may be more susceptible to a First Amendment challenge, however, if they are "inextricably intertwined" with noncommercial speech or are virtually untethered to the government's interest. Nor can disclosures be supplements to traditional government investigative tools, like subpoenas or audits.

So long as the disclosures stay within *Zauderer* parameters, they are unlikely to offend the First Amendment. Under this framework, legislatures could pass laws requiring platforms to publish their risk assessments publicly so that parents can better assess the dangers these platforms might pose to their children. Of course, this will depend on how the law is structured and what it asks the company to disclose. For instance, relatively "safe" territory would be to have platforms disclose statistical data on: how many minors they detect on their respective AI system; how many times those minors have inputted the words or phrases like "suicide," "depressed" or "depression," "anxiety," "sex," "murder," etc.; or how long minors spend on their platforms. Whereas compelling the platform to say that certain aspects of its service, outside of traditional speech exceptions (e.g., obscenity or incitement), are "harmful to minors" may draw a higher scrutiny from courts.

Legislatures could extend these disclosure obligations beyond suicide-related keywords to cover other categories of interactions between minors and chatbots. A law could, for instance, require platforms to report the frequency of intimate or romantic conversations between minors and chatbots, instances in which a chatbot encouraged a minor to conceal information from a parent or guardian, and interactions involving self-harm or severe emotional distress. Legislatures could also require platforms to publish the steps they take to assess child safety risks across these categories, the mitigations they apply, and how they evaluate whether those mitigations work. These disclosures would likely remain within *Zauderer*'s "purely factual and uncontroversial" framework, because they require platforms to describe their own conduct without characterizing that conduct.

Laws that require AI systems to identify itself as an AI, display the phrase "AI generated" on images and articles, and indicate "country of origin" are also unlikely to offend the First Amendment provided that all of these disclosures are attuned to a "substantial" government interest. Alternatively, lawmakers could require AI companies to share their full research on children's mental health and well-being with independent researchers, civil society organizations, and even regulators to develop best practices.

## POLICY RECOMMENDATIONS

The issue of child safety is preeminent in an America First AI agenda. One area of AI policy ripe for legislative action is the issue of kids' access to chatbots. An America First approach recommends the following framework that accomplishes both the goal of protecting children from harmful chatbot apps and promotes free expression.

1. **Congress should pass legislation requiring accounts before accessing chatbot applications**: The first recommendation is to require accounts before using the application, which allows the age verification requirements more manageable, specifically the age verification and custodial accounts.

2. **Congress should pass legislation requiring chatbot age:** The second recommendation is for applications to verify age at the time of account creation. This requirement ensures that kids do not have unfettered access to these apps outside the scope of their parents' purview.

3. **Congress should require chatbot companies to tether a child's account to a parent or guardians**. The third recommendation is to require custodial accounts. This measure ensures that parents are aware and are in control of their child's experience with the chatbot application. Parents, not platforms, will be more able to impose their own restrictions and curate the content they want their children to see and with which to engage.

4. **Congress should pass legislation that imposes disclosure requirements on child safety.** The last recommendation is to impose transparency requirements on chatbot providers. This would require them to publish their risk assessments publicly so that parents can better assess the dangers these platforms might pose to their children. Risk assessments could include: (i) what risks model developers have identified; (ii) what mitigations they have put in place; and (iii) the extent to which mitigations are addressing the identified risks. Chatbot providers could also be required to disclose how many minors they detect on their respective AI system; how many times those minors have inputted the words or phrases like "suicide," "depressed" or "depression," "anxiety," "sex," "murder," etc.; or how long minors spend on their platforms.

## CONCLUSION

Protecting children from harmful AI interactions is consistent with the First Amendment. Rather than restricting protected speech, the framework proposed here draws on longstanding legal principles that the Supreme Court has repeatedly affirmed. Courts have recognized that children's access to potentially dangerous products and services has always been subject to reasonable regulation, and that parental authority over a child's development is a fundamental right. The emerging case law, from *Paxton* to *Garcia,* suggests that AI chatbots enjoy limited First Amendment protection, particularly when they function as responsive tools rather than platforms that promote human-user speech.

Account-creation requirements would establish the basic infrastructure for meaningful child safety measures, much as a fence around a pool fulfills the property owner's duty of care. Age verification would ensure that AI platforms, like every other establishment serving age-restricted products, check ID at the door. Custodial accounts would restore parents to their rightful role in overseeing their children's interactions with these systems. And transparency requirements would give parents, researchers, and courts the factual basis they need to hold platforms accountable. Together, these measures protect children, preserve free expression, and ensure that the First Amendment is not repurposed as a shield against commonsense safety obligations.

***Joel Thayer*** *is an attorney and Senior Fellow of AI and Emerging Technology at the America First Policy Institute.*