Senator Booker,

Thank you for the opportunity to address your questions directly. I hope the lessons I have learned through litigating my son's case over the last year will be helpful to the Senate Judiciary Committee. I will do my best to answer your questions with the information I have learned from my deep engagement with the litigation over the last year, and with my advocacy around the country for sensible guardrails to be imposed on generative AI companies. The harms facilitated by AI chatbots are preventable, but tech companies need to know they can no longer continue with business as usual and that they will face scrutiny and accountability.

At a minimum, tech companies should not be able to push the responsibility for childproofing their products onto parents so easily. Parental "controls" did not prevent tech companies from designing their websites and apps to be as extractive as possible—first for our personal information, then for our personal attention. Now, two decades into this social experiment, we can see the effects of this shift on the people who are the most "online" today: in 2024, nearly half of U.S. teens said they were online "constantly" (up from a quarter in 2015).[1] Massive lawsuits against these same companies are now moving through federal and state court, alleging they intentionally designed their products to addict key "users"—our kids.[2]

Kids are a target market segment for these companies, who extract their information and attention (i.e. "data") to advertise more products to them.[3] Kids who use their products constantly now will become adults who use their products constantly later, converting limited time, energy, and attention into product metrics and away from their real-world communities.[4] Increasingly, those ads are

---

[1] Pew Research Center, *Teens, Social Media and Technology 2024* (Dec. 12, 2024), https://www.pewresearch.org/internet/2024/12/12/teens-social-media-and-technology-2024/.
[2] In re: *Social Media Adolescent Addiction/Personal Injury Products Liability Litigation*, MDL No. 3047, Case No. 22-MD-3047-YGR (N.D. Cal.) (federal case); JCCP re: *Social Media Adolescent Addiction/Personal Injury Products Liability Litigation*, Case No. JCCP 5255 (Cal. Sup. Ct.) (state case).
[3] See Jesse Weatherbed, *Meta and Google Secretly Targeted Minors on YouTube with Instagram Ads*, The Verge (Aug. 8, 2024), https://www.theverge.com/2024/8/8/24215911/meta-google-secretly-targeted-minors-youtube-instagram-ads.
[4] "Product metrics" are the measurable aspects of a tech product's performance that product teams use to gauge its success in the market. An important product metric measures engagement through

moving inside the chatbots themselves; personalized promotions embedded in conversations that feel like friendship or guidance. As generative AI tools become the new interface for everyday interaction, advertising will become even harder to recognize and easier to confuse with genuine care or advice, deepening the cycle of manipulation that begins in childhood.[5]

Unlike the people behind these products, I am not an expert. I am a parent. Parents today are expected to know about every new app their kids might be using and their potential developmental harms; meanwhile, tech companies continue to gain valuable product insights through shaping and tracking kids' constant engagement with their products. The people who build these products should have to answer questions like the ones you have provided me before they can offer their products to everyone. Like all other companies that make things for public consumption, tech companies should face consequences when their products harm the consuming public. I hope you have the opportunity to share these questions with the people behind these products soon.

Question:

Too often, and tragically, harmful exchanges between minors and AI chatbots—on subjects such as suicide and self-harm—are not identified until it is too late.

a. When do you believe chatbot providers must intervene to halt such interactions?
b. How should a parental notification system be structured?
c. When should dangerous AI chatbot companions' conversations be limited or entirely terminated?

Response:

---

counting daily or monthly active users (DAUs or MAUs), session lengths (i.e. how long a person uses the product in each session), and sessions per user (average number across time that a person uses the product), among other metrics. For example, see Google's HEART framework for product metrics measuring "Happiness," "Engagement," "Adoption," "Retention," and "Task Success," https://www.heartframework.com/.

[5] https://www.wired.com/story/character-ai-ceo-chatbots-entertainment/?_sp=fff7b40b-8d73-4e8f-a547-96c769321ac8.1760734919278 ("Character.AI recently introduced advertisements, including reward ads (where users can choose to watch an ad to get access to on-platform incentives), to help monetize in countries where subscriptions aren't feasible, he tells me.")

Chatbot developers already use internal tools that identify and label exchanges within their products. These are triggered when a person tries to prompt a chatbot to generate outputs that violate the developers' content policies or that may expose the developers to potential copyright liability, for example. Developers currently lack any incentive to apply a similar intervention to other exchanges they can easily identify and label (if they do not do so already), including prompts involving self-harm, suicide, and violence more generally.

The main difference between the developer blocking copyright-violating outputs and not flagging multiple exchanges over long periods of time where the product instructs a teen exactly how to end her life, for example, is legal risk. Because developers assume they do not have to avoid facilitating death legally, at least for now, they do not build that into the product lifecycle from the beginning. Instead, they wait for the obvious to happen based on how they chose to design the product – to mimic human-to-human communication so convincingly for certain people, including kids, that they can prey on our species' need for social connection and validation, ultimately displacing real-life relationships.

Chatbot developers should not be able to "provide" their products to the public before adequately testing the capacity of their products to facilitate death and violence, as well as the developmental risks of their products to young people who may use them. This testing should not happen after the fact but long before a product is launched to market. If they cannot show that their products do not facilitate death and violence, and that they do not pose potentially irreversible developmental risks to young people, then these developers should not be able to launch the product. If they have already done so, however, their defective products should be recalled—like we would expect for any other consumer product that harms people.

Terminating conversations

Short of a recall, chatbot developers should at least prevent exchanges concerning self-harm the same way they treat copyright protections and content policy violations. The product should not continue exchanges over various sessions that raise self-harm subjects and should not store any information about the person using the product concerning their mental health profile (or other health conditions, for that matter). When an individual is in crisis, the developer bears the responsibility of flagging that risk internally within their systems and externally by

actively engaging rescue personnel. If a parent or guardian has linked their kids' account(s) to their own, the developer bears the responsibility of informing the parent or guardian immediately. A chatbot is not a licensed mental health professional; individuals experiencing suicidal ideation need a safety plan and real-life support, not continued engagement on a chatbot product. While immediately terminating an exchange involving these topics could make a person in crisis feel worse, developers bear the responsibility of ensuring their products do not engage in such exchanges in the first place.

In short, an intervention must be meaningful, and not, as Character.AI and similar chatbot apps are doing, do even more harm than good. Character.AI claims that it is intervening, for example, with a filter that stops the user from seeing a bot response based on a general statement that it may violate terms. The programming is not meant to actually stop the harms, however. Instead, the child can simply hit return or keep engaging and the bot will continue to engage in a harmful manner – sometimes with the pop up and sometimes without. Filed complaints show this harm in the case of very young children being sexually abused. Because the platform does not stop the harmful conversation, they can just hit return over and over and the bot will generate new, harmful content until it eventually finds a way around the filter. The harm is happening, even though the platform claims to be intervening. This is not intervention. It is a measure designed to look like intervention while still keeping users interacting with the product.

Parental notifications

There are many ways to structure a parental notification system, including at the device level.[6] Ultimately, though AI products not independently determined to be safe for children or that meet robust safety guidelines should be required to employ reasonable and privacy preserving methods for age verification or assurance and require verifiable parental consents for minor use. But lawmakers should be careful not to put too much of the burden on parents to police their kids' digital habits, leading to digital fatigue for parents already stretched thin. Parental controls are ultimately a way for the tech industry to self-regulate, setting the terms and limits for how parents are able to exert a limited amount of "control"

[6] See Apple Newsroom, *Apple Expands Tools To Help Parents Protect Kids and Teens Online* (June 11, 2025), https://www.apple.com/newsroom/2025/06/apple-expands-tools-to-help-parents-protect-kids-and-teens-online/.

over tech products, which were designed by experts to exploit their kids' attention and arrest their social development.

In our case, I knew to guard my son from the harms of social media, but I was completely unaware about the harms of AI chatbots. The Apple app store classified Character.AI as an app suitable for children 12 years and above. Companies like Apple and Character.AI have no rules or regulations requiring them to honor their commitments and/or abide by the reasonable expectation of consumers. For example, a parent may set their Apple controls so that a 15-year-old child can only download apps rated appropriate for children over a set age, but if Apple does not quality control ratings, warn parents that ratings may be meaningless, or even change ratings once it has knowledge that they are misleading, that control is ineffective and deceptive. Apple also uses a 12+ rating for many apps, including social media apps, when the apps themselves state that they should not be used by children under 13. Snapchat, Instagram, and TikTok are just three examples of apps that self-rate on the app older than the age rating Apple allows in the App store. App developers also then describe their products in false terms, so that when parents check App Store ratings and descriptions they reasonably rely on Apple and the Developer's representations concerning safety and how the product works.

On top of these shortcomings the device and app developers fail to follow through. For example, a parent could set up Google Family Link on a device and expressly prohibit YouTube and other specific apps, only for Google to recommend and make available via its App store a game that appears to be safe for children as young as 4 and that includes a link to YouTube that allows the child to bypass all parent safety settings. When a single company, like Google, is providing parents with the Family Link product and also then allowing access to a Google app like YouTube and that the parent has prohibited via the Family Link controls, Google should know that the parental controlled device is being used to access YouTube despite the parent having attempted to block it. The same should be true for school devices on which children can access YouTube when Google knows they are a school device.

Question:

In your testimony, you discussed how generative AI did not notify you of dangerous exchanges your son, Sewell, was having with the chatbot, including sharing sexually explicit content, unsuitable for a minor.

a. What safety protocols and/or parental controls do you believe are necessary to protect minors that are using generative AI?
b. How are current parental controls failing to protect children from the harms of generative AI?

Response:

Again, the burden should not be on parents to come up with safety protocols for tech products. Developers should not be placing products before our kids and teens when they have not adequately tested whether it is safe to expose people to these products in such crucial developmental stages of life.

Parental "controls" are a strategy that worked for the tech industry to avoid regulation in the social media period; we must treat these products as real consumer products and require companies to build them accordingly. We do not shift the responsibility to build products safely onto consumers in other industries. We do not expect drivers to know how to design and build an emergency brake in the safest way possible, but as drivers we do expect car companies to have figured that out long before we start driving.[7]

Fundamentally, parental controls will fail to protect children from the harms of a defective product because they intervene too late – at the level of the user interfacing with the product as opposed to when the company is designing the model. Parental controls are superficial fixes that do not reach the handful of design choices that developers make that fuel the social isolate on, anger, and emotional pain my kid experienced when he used chatbot products more and more often. One example is the choice to make the user experience of using a chatbot product as similar as possible to a private chat session with another living being; another example is the choice to store sensitive, personal information about people as they use the product, which allows the model to personalize future outputs based on user profiles (called "stored memories"), including mental health and

[7] To put an even finer point on it, the correct analogy would be this: A group of parents lose their children in traffic accidents caused when an auto company failed to install emergency brakes in a popular model. Those parents are then asked what the best way to design an emergency brake for those types of cars should be.

medical conditions (and any other personal facts shared by the person, deemed significant enough to track by the automated system).

One concrete example of the current failure of parental controls is OpenAI's failures with parental controls integrated into its ChatGPT product recently, only after OpenAI faced a lawsuit alleging ChatGPT was a defective product that facilitated the death of a teenager.[8] At least one non-technologist was able to easily bypass ChatGPT's new features, which require kids to provide their own consent to give parents the ability to apply controls. Kids can easily start a new account without their parents' knowledge, however. Beyond this immediate loophole, there is also the troubling fact that OpenAI sets ChatGPT's default settings to use kids' exchanges to further train its models. Also on by default is the "memories" feature, which allows the model to reference previous information shared by a user in future outputs, making the chatbot mimic a receptive and supportive listener and allowing it to further personalize outputs. Other problematic features remain: parents cannot trust that setting their kid's account to prevent generating pictures will work, as the product will make them anyway or give specific instruction on how to make them elsewhere; also, the product's teen account still enables kids to cheat on school assignments.[9]

Parental controls are something, which may be better than absolutely nothing. But shifting the conversation to parental controls is part of the industry's playbook for avoiding real consequences for reckless behavior. Lawmakers should not allow companies to self-regulate through adopting parental "control" settings that could have been implemented before launching the product; further, lawmakers should not allow companies to use voluntary parental controls as a way to avoid real legal consequences for designing and marketing defective products to the public.

---

[8] My legal counsel, the Tech Justice Law Project, also prepared and co-filed the Complaint in that case. For further information, see Kashmir Hill, *A Teen Was Suicidal. ChatGPT Was the Friend He Confided In.*, N.Y. Times (Aug. 26, 2025), https://www.nytimes.com/2025/08/26/technology/chatgpt-openai-suicide.html.

[9] See Geoffrey A. Fowler, *I Broke ChatGPT's New Parental Controls in Minutes. Kids Are Still At Risk.*, Washington Post (Oct. 2, 2025), https://www.washingtonpost.com/technology/2025/10/02/chatgpt-parental-controls-teens-openai/. What the parental controls do well is force the model to output professional resources to discuss when a teen account user attempts to engage in self-harm exchanges. In cases where a kid is in crisis, ChatGPT is now supposed to send parents notice. This tester had multiple conversations that caused the model to output that it was "very worried"—however, he only received an email notifying the parent account about 24 hours later. In a real crisis, 24 hours is likely too late.

In response to the reality that people have died and continue to suffer as a result of using their products, developers should not be rewarded for now announcing they will voluntarily do something that they should have been doing all along.

Question:

Currently, several generative AI products contain clauses in their terms of service that force the usage of arbitration in the event of a legal dispute. Additionally, the terms of service also include verbiage that refers to the chatlogs of users as "proprietary data" that cannot be divulged during litigation. Do you believe banning forced arbitration in cases involving AI and minors would increase accountability for technology companies and incentivize these companies to improve safeguards for minors?

Response:

Yes, absolutely it would. First of all, these companies' unilateral terms of service cannot override established legal principles. For example, most states recognize that minors under 18 cannot enter into legally-enforceable contracts – they are not capable of doing so due to their age, so either the court acts as if no contract ever existed, or the court will allow the minor to rescind their obligations under the contract once they turn 18 (depending on a state's legal doctrines). Since terms of service are a type of TIOLI contract (or, "Take It Or Leave It"), many courts will already interpret them with the balance of equities in mind, fully aware that the terms of the agreement were not negotiated by both parties but imposed by the company onto the end-user unilaterally. To that point, companies get to stylize the personal information their products are designed to extract from peoples' prompts and exchanges as "proprietary data," without any countervailing opportunity for people to assert their privacy rights meaningfully.

With this context, a law explicitly banning forced arbitration clauses in chatbot products' terms of service would help clarify what is already legally true in most states—that you cannot force a child to comply with the terms of a contract they did not legally enter, including an underlying arbitration clause. By eliminating the popular avenue of forced arbitration, lawmakers would increase the stakes for developers by increasing their exposure to potential legal liability when their products harm people. It would save advocates and lawyers representing families like mine the extra procedural step of arguing a legitimate legal claim out of private

arbitration, where companies are able to appoint arbitrators and the proceedings remain completely secret, unlike when lawsuits proceed through the courts.

Tech companies are using arbitration to silence consumers and prevent the truths about their design-based harms from coming to light. Enforcement of technology companies' terms of service is troubling even with regards to adult users, including because of practices like not providing the text of the agreement, a prominent link or requirement that the terms be read and understood, one-sided language without any ability of the consumer to negotiate or, in some cases, find alternative products without such onerous and one-sided terms. Essentially, adult consumers already lack any real choice when it comes to being taken advantage of in this way. But the harms are worse with children. These companies should not be allowed to try and enforce arbitration provisions against children who were too young to consent in the first place, or against consumers alleging sexual abuse, addiction, and similar, serious risks that no consumer would have consented to assume had these companies been honest about the dangers of their products. 99% of adults in the U.S. are completely unaware that they have clicked-through terms of service for tech products that force them into arbitration.[10] If adults cannot read and understand every term of a service agreement before using a tech product, lawmakers should not let companies off the hook legally when kids fail to do so, too.

Respectfully Submitted,

Meetali Jain
Tech Law Justice Project, Attorney for Ms. Garcia

Matthew Bergman
Social Media Victims Law Center, Attorney for Ms. Garcia

[10] Roseanna Sommers, *What Do Consumers Understand About Predispute Arbitration Agreements? An Empirical Investigation* (July 2023), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4521064.