Tue, Sep 22, 2026 12 PM Rayburn House Office Building

SUMMARY KEYWORDS

AI speech, First Amendment, chatbot output, legal liability, user rights, listener rights, speaker rights, editorial discretion, training data, content moderation, government coercion, jawboning, tort law, cyber security, AI agents

SPEAKERS

- Corbin **Barthold**, Senior Litigation Fellow, NetChoice
- Kate **Ruane**, Director, Free Expression, Center for Democracy & Technology
- Megan **Iorio**, Senior Counsel, Electronic Privacy Information Center (EPIC)
- Cristiano **Lima-Strong**, Senior AI & Government Reporter, Bloomberg Government

Cristiano Lima-Strong 01:55

Thank you all for being here today. As Tim alluded to, you know, this is a very timely conversation we're going to have today. We've all seen there's been an explosion of interest around AI and concerns around risks and harms in the space in recent weeks. And as we see policymakers begin to formulate response to that, more of a response than we've seen, there's a possibility that that could collide with some of the discussions that have been going on for years now in legal circles around AI and to what extent, if any, First Amendment protections and principles should apply in this space. And so we're going to be, and of course, there's also parallels to that and some of the discussion that's been unfolding both here in Washington and also in courtrooms around the country around social media regulation, and some of the lines there. We're going to dig into some of the potential parallels there, where they diverge, as well as where things could be headed with this latest wave of AI developments. I wanted to start off the conversation though by reading a couple things from some of the most influential thinkers I've been exposed to, first being the founding fathers. We're going to be talking about the First Amendment a lot today, so I figured I'd just read it verbatim to get us going. Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof, or abridging the freedom of speech, or of the press, or the right of the people peaceably to assemble and

to petition the government for a redress of grievances. The second item I'd like to read is a text that I got from my wife when I told her I'd be moderating a panel on AI and free speech. Her response was, "But it's not people? Question mark. I'm sharing this not because I'm trying to get in trouble later, but I think this is probably a common response that you get if you ask someone about First Amendment principles and how they apply to AI. So I figured we could start sort of at a broad level, and I'd love to hear from you all. You know, as Tim was talking about, these are machines. This is code. Like, why are we sitting here talking about the First Amendment? So maybe we could just go down the line and start, Megan.

Megan Iorio 04:02
Yeah, I mean, there's lots of reasons why we'd be talking about the First Amendment. I think the the most obvious reason is that the the AI that we have that's generating speech is is is providing people with Information and sometimes with particular viewpoints. And whenever you have a technology like this or a media like this, you have people in government who want to control what information people are exposed to and what viewpoints they're exposed to. And this is probably like the largest First Amendment type of danger that there is in this context and comes closest to like the core of what the First Amendment is supposed to protect against, but then there are other aspects of First Amendment doctrine and and speech interests that apply in this context. But it's complicated. It's very complicated, and there's no. Like there's no extreme answer. It's not that it's not speech, and or or it's not that everything is speech or everything's not speech. Some of the things that developers do to create chatbots and other AI systems are expressive, but a lot of them are not, and the the whole difficulty is going to be in figuring out which aspects are expressive and which are not, and how the government could intervene to protect people from harm.

Corbin Barthold 05:36
Great intro question and great first answer. I mean, yes, an LLM is not a person. A video camera is not a person. A video camera is a tool. An LLM is a tool. Both are tools of speech. And as Megan was just saying, that means that the First Amendment comes immediately into play because I think we naturally, and for good reason, when we think of the First Amendment, we think of speakers. But the First Amendment talks in terms of speech, and so first of all, attacks on LLMs that try to censor what they can say-you should think of them first and foremost, as attacks on you, there's a lot of First Amendment case law for the notion that you have rights as a listener to receive information, to make of that information what you will, to learn from it, to quarrel with it, to judge it. So there are several cases that hold that you have a right to information that has not come from a speaker who him or herself has First Amendment rights. Once you are dead, you may not have First Amendment rights, especially if you're, say, Aristotle and you died in Greece 2,000 years ago. But I have a First Amendment right to read what Aristotle said. So you have listener rights. Then you have a right to use tools that implicate speech. So if state says, "Well, we're not going after what the newspaper can say. We're just going to tax the ink and the paper that it uses, that is a First Amendment problem that violates the First Amendment. So there are strong First Amendment issues here. I actually very deliberately bring it up third, but the AI firms also have First Amendment rights in how they craft a speech service. There's a lot of expressive choices that go into how an LLM is designed, the data that goes into the LLM. Tons of stuff happens during reinforcement learning, right down to does the LLM have a sense of humor? How strongly does it push back on you when it thinks that you've said something wrong? We had that issue not long ago about sycophancy and people thinking that LLMs are too agreeable. How you calibrate that is an

expressive decision, and the Supreme Court has said multiple times that the government cannot be in the business of deciding what speech is valuable versus not valuable, or bad speech, or virtuous speech. So even just basic decisions about how information is presented to you are First Amendment protected. So up and down the line, there's a lot of First Amendment issues at play here, particularly with chatbots. Although I agree, things get quite complicated as we go into the brave new world of agents and all kinds of stuff we can dive into. But at least for me personally, when I look at the chatbot, and you're right to use it to learn and increase the marketplace of ideas and go out and speak with what you've learned from it, I find that at least pretty straightforward. Kate,

Cristiano Lima-Strong 08:55
Kate, thoughts?

Kate Ruane 08:56
I I don't have much to add on top of what Corbin just said. I would just say I agree. Robots don't have First Amendment rights, but as Corbin pointed out, the the user First Amendment rights to both access and receive information, and to and to actively develop the information that it gets that you get from the chatbot, those are significant First Amendment interests that are in conversation and mutually reinforcing with the First Amendment rights of the developers and deployers of these sort of speech systems. And then the last thing I would say to reinforce those things is like the First Amendment has never been absolute. If we determine that the First Amendment does apply to aspects of the design, deployment, and then use of AI systems-these systems that are generating new text, generating new images, creating and distributing information, accessing information, and arranging. And presenting it to people, if we decide that there are speech interests involved here that the First Amendment protects, and I think that there are, the First Amendment has also always had limitations, and it is in that conversation where I think the rubber is going to meet the road eventually, where we're going to have to have some complicated discussions about what can be restricted and when, what are what are the ways in which a chatbot leading a chatbot providing certain kinds of information leading to harm in the world could potentially lead to liability and for whom? Like who is responsible in that chain is it the person who made the chatbot? Is the person who deployed it? Because those are two different people sometimes. Is it the person who designed the safety policies that sometimes also a third person in this conversation because because there are companies that provide third party that third party services that develop that can be plugged in to some of these to some of these systems, in what ways do existing civil rights laws apply? In what ways do they not? Do they need to be adjusted? And and and then, as Corbin pointed out, what about when the agent can do something? What about when the AI can go and go out in the world and take a real-world impact. What happens when it buys your airline tickets, but it does it wrong? What does the First Amendment have to say, if anything, about that? Those are those I think are the open questions. But whether there are speech interests at play, I think it's pretty clear that there are, and they are they particularly belong to the users who are seeking to understand their world now through this new mediator.

Cristiano Lima-Strong 11:48
So I referenced that text message. One of the reasons there was kind of an inferred speaker there in the question, and that's something some of you touched on. And in terms of some of these different distinctions between speaker rights, listener rights, and what what sort of might be considered? Who are we talking about when we say the speaker when it comes to to AI and the First Amendment?

Megan Iorio 12:08
That's sometimes the question. If folks have heard of the litigation in Florida over character AI, in which a child died by suicide after having extensive conversations with a chatbot that took on the character of a Game of Thrones character. In in that case, the the companies there made the argument that it actually didn't matter, and they didn't need to identify any kind of speaker for the the chatbot outputs. All that mattered was that someone was listening, and it was their right to listen, which, by the way, was a child that died by suicide, because in part, possibly we never got to that stage. What was outputted and the lack of safety measures in the chatbot, and so, really, the question is, I think, and that historically, the First Amendment has protected the rights of people to speak and to listen, and so in all of First Amendment doctrine, even ones involving corporate speech, it's the people's right to say something and people's right to receive that information that makes something speech. And then we get to the other question of who has rights that to assert. Whether a speaker has a right to assert in that case, if they're in a foreign country, they don't. But if they're here, they do. But if the listener is here, they can assert the right here. The question I think in these cases, and we're going to differ on this: Is who is the human who is responsible for this speech or expression? Is there a human at the end making decisions that result in the expression that is being regulated? And if there are not, and if the company refuses to say, yeah, that's our speech. I think that that shows that there is something fundamentally off with the First Amendment arguments. Well, again, we're going to differ here.

Corbin Barthold 14:32
We we may agree more than we think. Although I will start with the hard version of the listener right that people, I think, need to remember when we talk about speech that we get upset about speech and we argue over speech precisely because speech is powerful. Speech can cause emotional harm to people. Speech can change the world. It can start revolutions, and that's precisely why we protect it. So. Great Supreme Court case, Lamont versus Postmaster General, which I love to bring up in all kinds of contexts. Corliss Lamont was this kind of a crank, and he wanted to receive an issue of the Peking Review, which was this piece of Maoist propaganda. And the law said that you had to check a box on a card and submit it to the post office and say, no, no. I really do want to receive that piece of communist propaganda, and the court struck down that requirement because, as the decision said, you, as a freeborn American, have a right to receive what the government thinks contains the seeds of treason, because we treat you as a citizen to think for yourself and not as a subject who's subject to you know paternalistic constraints, so you have a listener right even when speech could potentially be harmful. Now, having said that, I do agree that the LLM in speaking doesn't have intent, and it sounds like it has intent. It talks like it has intent, but it doesn't, and that creates this weird problem with the First Amendment because there are various forms of unprotected speech under the First Amendment, such as true threats or incitement to violence or criminal conspiracy, where an element of that crime is intent, and I agree that we're probably in an intuitively bad place and a legally bad place. If the rule is that because LLMs don't have intent, it's just anything goes and they are not accountable. I think ultimately you do need to have some kind of principal-agent relationship where the creator of an LLM is responsible for what that LLM says, no different than if a corporation has an employee and that employee goes rogue and commits a crime. If it's within the principal-agent relationship, the corporation has a problem. The one caution I would give is I think there is a belief or an effort to start from first principles and and or a blank slate, I should say, and act like we can then impose rules on an LLM that would not be imposed on a normal speaker, and if the LLM is held to the standard that if it sounds like

it's inciting violence, we should treat it like it's inciting violence. I think that's probably not a bad place to end up, but some of these very tragic suicide cases-not-we'll have to see what happens with the logs and the chats-and I have no idea what could end up happening. But some of them, people are trying to hold the LLM to a much higher standard than a human would be held to, and I worry about that bleeding into the wider speech environment, where, for example, you're at risk of liability for someone's suicide simply because you talked about them with them about their emotional troubles, or you're potentially liable for a shooting because you talked to a person about guns, and you didn't actually show the kind of intent that we normally require under the First Amendment. So we need to be careful. It's a new terrain as we set the new rules. I do hope that we keep the normal and traditional First Amendment principles squarely in mind.

Cristiano Lima-Strong 18:17
So we've referenced a couple instances of AI chatbot outputs, and we were talking before this panel that you know a lot of the conversation a couple years ago was very much focused on that. Recent weeks, we've been talking about sort of rogue agents and what happens when humans lose control of AI. But I want to I want to go back to sort of the the chatbot output scenario. So some some of the discussions that we saw in Washington is what happens if and if someone feels they've been defamed by an AI output or if it's led to discrimination or misinformed. How how are you guys thinking about First Amendment principles and how they apply in instances like that to an output of a chatbot? Kate, do you want to take that one?

Kate Ruane 19:01
I sure. I'll. I can start with that. I guess the the way I've started to think about it, and again, like this is these are all very open questions, is one of the things that the First Amendment doesn't account for or do a super great job at is uncertainty. Chatbots themselves are producing new speech that is the result of guidance, but also the also interactions with users. Its outputs are unpredictable, even when it has rules that say you cannot say that. Like you, even even if it has rules that say you may not lie about particular individuals, one those those rules can be overwritten by by determined users. Two, they can be I won't say circumvented, but they can be they can be essentially broken. If certain things happen within the the creation of the output, where the for lack of a better word, the the output is inaccurate for one reason or another. So sometimes it at least once in the past a chatbot has been accused of defamation because it confused two people with the same name, so it was being asked about one person who was a journalist in in the United States, but it produced information that conflated that person with another person with the same name who was in jail for child molestation, and so it was telling people that this person, who was actually two separate people, was a journalist who committed child molestation. That's incorrect, and if uttered by a human, would be defamatory. And and so you know there are there are circumstances where chatbots have done this. The question to me is like how how do we how do we go back in the chain to to where the to where the locus of responsibility is, and how do we determine what the what the responsibility of the humans in charge of the creation of that service is, and did they perform enough to try to safeguard against the mistakes that the that the chatbot made to determine whether liability should attach? Those are the questions that I that I would want to ask in terms of how liability should be structured when chatbots do things that are or would be the source of liability potentially if a human did it, because the system itself is is a is an uncertain output creator, and if you cannot have certainty in what outputs are going to be, then you have to have some level of uncertainty and some type of and some type of new standard back where the back where the

chatbot is being created and then deployed. That that's sort of how I am trying to think about it, but I I am open to to conversation and pushback and other and other discussion because I do think that this is new terrain that we are that we are traversing here.

Corbin Barthold 22:09
I too am an uncertain output creator, as are most humans. The to to just it builds a bit on my last answer because I you asked about who's the speaker, and I was talking entirely in the environment of an individual talking to a chatbot. And I think that clearly we need to have some kind of responsibility for the creator of the LLM. Once people take speech and move it out into the world, you do have some responsibility as a speaker yourself, you can learn somebody can tell you something, and it's bad information, and you go and repeat it in public, and you can still get in trouble for that. And I don't think that's any different with a person or an LLM. Defamation is an interesting case because normally, if you say something in private, you're not going to have a very good defamation case because a whole element of defamation is that a lot of people heard it, and that's your damage. So you would potentially want to pin responsibility on the person who just takes that without some healthy skepticism and turns around and repeats it in the public to the public. Obviously, if you have an AI and you've created it, and you've let that AI go out on into the world, and it's speaking in public. Then you've taken on potential risk of of what it might say. But just generally, I would note that while AI is revolutionary and it's new, in some ways there's nothing new under the sun. I don't know exactly what I'm going to say to you guys. I have beliefs in my head, and what I say is probably going to track that. You could call that my training data. I'm going to make choices about what I say. I'm going to do my best. I'm going to try not to blurt out something defamatory. There's some edge case where that's a possible thing that could happen, and so I don't know if the rules necessarily need to be drastically different for AI versus humans in that context.

Kate Ruane 24:09
And I wasn't trying to suggest that they aren't, but if you say something defamatory now, we know that that goes back to NetChoice currently employs you. Let's just put it back to me

Corbin Barthold 24:20
for the moment. Well,

Kate Ruane 24:21
maybe. Well, theoretically, if you were like, if you were a chatbot making making those noises that wound up being untrue statements that harm someone else's reputation, and they were published, and they were published to someone other than the person who, to whom they applied, right? That's that's enough to get you defamation. It might not be enough to get you damages, but it's at least enough to get you to get you defamation. The robot itself doesn't necessarily have liability. You can't sue it in court, so you would have to sue somebody, and the the question is who is that somebody? At which point, in which point in the in the development and deployment chain, does that have to happen? Like I think that's how I think that's most of how I'm I'm trying to think about that. Yeah, no,

Corbin Barthold 25:11
that's fair. And I think at the moment we have a pretty neat divide between chatbots that talk to individuals in private, and then people who go out and speak in the world based on what they've

learned with agents, but that or from AI, and but as you're alluding to, we're quickly getting to the situation where if you have some agent that's a fifth generation and nobody knows who owns it, yeah, that's a whole new.

Kate Ruane 25:31
Or even just if or even just if I if like if I'm using the chatbot and it tells me something untrue about you, that could be enough of a publication under defamation law, so it is it is happening in these circumstances. I have I have I have very little doubt that it is happening right now, but I take your point that the the reason we're not seeing a ton of it is because these are happening on one-on-one interactions as much as it doesn't much as it does between people, right? The problem right now is something defamatory at any point. Yeah, right. Right now, my my damages

Corbin Barthold 26:06
are just that Kate, you think less of me. Yeah, yeah, yeah. Well,

Megan Iorio 26:09
I was just going to say, like, this conversation is more about how tort liability applies in this context than like the First Amendment and whether it's speech. But I think if there is some like parallels here, and Kate was getting at this, wherein the question here, as I think the question also is in the First Amendment context, is what do the humans have control over, and it is the processes that they put in place from like selecting training data to fine tuning to to to training et cetera et cetera all those different parts and then like the safety measures that they don't do or do not put in place that are the things that the people are responsible for that they could potentially be held liable for under tort law but we need to kind of flesh out what those tort duties would look like because this is a very different context, and it's going to be similar, I think, in the First Amendment context as well. When we're looking at what all these levels the companies are actually doing, and if the if the government intervenes and requires them to do something else, what is the impact?

Cristiano Lima-Strong 27:21
Yeah, also I'm sure we could do a separate panel entirely on on the liability question. But one thing I wanted to to circle back on that a couple of you have mentioned, but we haven't dug into as much, this idea of whether something's expressive, and I think that that's a key question here. There was a very famous paper in 2021 by Emily Bender and Timnit Gebru, who argued that LLMs are basically stochastic parrots. Basically, this idea that it's like all this data that's being analyzed, statistics and probability, and it's being used to produce or mimic responses Parrot, something that is not akin to comprehension, and so you could maybe and please correct me if I'm I'm getting the legal train of thought here wrong, but you could extend this to say that this is not expressive. So you know, Corbin, you were talking about oh I'm I'm a random output generator, but there's a distinction some folks would argue quite random.

Cristiano Lima-Strong 28:23
But so how do we think about the the concept of whether something is expressive, in terms of how we consider whether something should be protected under the First Amendment when it comes to AI? Whoever wants that.

Kate Ruane 28:37
Oh no, Megan, please please go.

Megan Iorio 28:40
I actually think that there's a few different levels to this question, and I actually think that the CDT paper that you all put out did a good job of like highlighting the different ways that speech interests come into play in various systems, even when the thing that we're talking about is not expressive, so like we can say chatbots are not expressive at all just because of all the reasons that were said here today. But if the government were to come in and impose some kind of viewpoint restriction on the outputs of a chatbot, it doesn't matter that the chatbot outputs are not speech. They're not expressive. The viewpoint limitation would impose a speech restriction on the chatbot and on the company. Whoever was is is is putting out the chatbot. Because, for instance, let's say no woke chatbots. That requires the developers to remove certain information to prevent it from generating certain information, and that is the kind of intervention that interferes with expression, even when the. Like the object might not otherwise be expressive. This is complicated, but like this is a line of First Amendment doctrine. The same thing happens when if the government were to come in and say you can't have like these ideas in in a chatbot, and the same is like we have a whole line of precedent about must carry, which is the part of like compelled speech and editorial discretion. All of those, all of those, that line of cases is part of this doctrine, wherein if the government requires someone to carry some idea or to not carry something, and the opposite, that is an interference with that person's ability to decide whether to carry it or not carry it, and that is like a recognized as Expressive, and you don't really need to do much more analysis on that end, and and that's where we get to like the one of the more recent Supreme Court decisions, Moody v. NetChoice, wherein Texas and Florida both passed these laws saying that social media companies had to carry all messages coming in on certain topics or from certain people, and that conflicted with the social media companies' policies to remove certain types of speech. That is, the social media companies did not want to carry that kind of speech, and the Supreme Court. This case got up to the Supreme Court, and the the the court said, "Look, this is this is a very clear like compelled speech must carry editorial discretion case, wherein the the social media companies will remove this these messages if not for this government intervention and requiring them to carry these messages interferes with their expression. Now, in order to apply this, the and then the Supreme Court did say, though, that there are certain aspects of like a social media feed, the generation of an algorithmic feed that might not be expressive, and they for they use as an example a algorithm that the government were to. Like burden a company's ability to provide an an algorithm that just used user data to give users what they wanted, then that might not be expressive. And then we had a few concurrences, most notably Justice Barrett, who wrote an addition about this human element of things, like when we use LLMs to decide what content to take down, there might be a point at which that decision is no longer traceable to a human, but on the expressive end of things, in the lower courts, this has trickled down to courts looking at okay, what decisions of the social media companies in making algorithmic feeds are expressive and what are not, and we have a court out in California, a district court in California that's considering a law that limits social media companies from using certain surveillance data to generate algorithmic feeds, and that court noted that the companies in using this data don't seem to be doing anything that resembles any any of the decisions in these precedent, including the fact that they are not considering the content of the of the posts that they are selecting based on user data. It's all about analyzing the user and blindly choosing content that way, and that there it has no expressive aspect to it.

Megan Iorio 34:35
In this, in the line of editorial discretion cases, for instance, this case about parade organizers in Boston who put on an Irish parade celebrating Irish heritage on on St. Patrick's Day, the the court in that case had noted that yeah they might not. Have a particularized message in that case, but that they had a theme, and even the social media companies don't have even a theme when they are picking content along those lines, and so what courts are doing is looking at the specific mechanism that the government is intervening on, and asking, does it resemble in any way the kinds of communicative? Does it have any of the communicative aspects that, like, are in the past precedent? Sorry. Can

Kate Ruane 35:41
can I can I follow up with like an an example of what the what Megan's kind of talking about here? the The question goes back to the stochastic parrots article, I and whether chatbot outputs are expressive. I don't think those two things are necessarily opposed to one another, because it's not about whether the output itself is an expression of any single entity. Instead, the question is like, where are the expressive where are the expressive rights throughout the process? And the example I want to use here is training data. So the first editorial decision that a chatbot developer is going to use is what is the data that I am going to gather to train my LLM on, and it will make choice. They will make choices about what they want the eventual outputs of the system to be or not be, so they will do things like they will remove child sexual abuse material in order to try to reduce the reduce the likelihood that such content will be produced by an image generator, or they will go even further and have and some have tried this where they will remove things like the concept of a child from the training data, and then when you do something like that, you see in the ultimate output that other things are affected as well. So an LLM's a chatbot's ability to produce things like playgrounds then is is is changed by changes to the training data. If you try to remove the concept of nudity, it it can then make it more difficult for certain for other types of images to be developed because it turns out that if the chatbot understands, or if the if the AI generated system understands a body better, it can generate things like someone running in a more accurate way if it has a concept of what the human body looks like underneath clothes, but AI like AI companies are making these editorial decisions at the at the training data stage, and then at the and at the training and then at the pre-training stage, and then again in the input and output filter stage to try to determine what the chatbot should or shouldn't say, even though they can't ultimately predict the specific output at the end of the day. Each specific output, because you're going to get a different you're going to get a different output for the same prompt if it's if it's prompted at different times by different people and different users. That's just they know that's how it is, but they make editorial choices throughout the process that that that may not produce specific determined speech. But under the First Amendment, under Hurley under the Hurley case that Megan cited earlier, you don't need to know. You don't have to. You don't need to know exactly what the ultimate output is going to be. You just need to exercise editorial control over it in some way, and have a and have and and be trying to produce a particular kind of message or eliminate a particular kind of message like child sexual abuse material from your ultimate outputs. So, is the is the output itself expressive or not? Is I think a formalistic way of of addressing the issue. Rather, I would look at the process of the creation of it, and add in the fact that users are influencing that process as well by designing by designing specific prompts, by providing specific inputs to try to get the information that they want or need out of a particular system. So that's sort of how I think about the expressive question, like who's expressing what. It's not it's not really to me about whether the bot is saying anything. It's about everybody in the in the development and

deployment process, and then all of the users who are interacting with each other to create whatever the ultimate result is.

Corbin Barthold 39:51
Dive in real quick there. Maybe I come at it from a slightly different angle. So, as Megan noted, there have been some cases in the lower court. That have latched on to this notion of a specific human expressive choice, and I'm not actually sure that's right, or at least I think that could easily be taken too far beyond what the law currently says. I think it it both maybe gives a little bit too much credit to humans and creates a danger of limiting expression unduly. So, as far as I'm concerned, you mentioned ideas, and as far as if ideas are present, I would say expression is present. And actually, I would argue until maybe Justice Moody's concurrence in in Moody, that's not really something.

Kate Ruane 40:40
Justice Barrett.

Corbin Barthold 40:41
Sorry, pardon me. Justice Barrett's concurrence in Moody-that's not something we hinged a whole lot on. If you are running a library, maybe you just want to have a lot of books. You want to have a lot of human knowledge, but we would look askance under the First Amendment if the government started telling you you had to keep specific books out, even if you personally have kind of no idea exactly what views are in that library, and you know you could run a newspaper, and you're the publisher of the newspaper, and you just tell the editor, "Look, I own the newspaper. I want you to sell papers. Run whatever sells papers. You don't lose your First Amendment right just because you're seeking to maximize reader engagement. If you decide to run the the National Enquirer and you're running stories on the Bat Boy in a cave, okay, maybe that's an expressive choice. It's a really silly one, a bad one. You don't lose your First Amendment right. You know, Jackson Pollock could splatter the paint all over and not know where it's going to land, and he still has an expressive right. So I totally acknowledge that it could be that the rule ends up being we we put a human into the loop, and that's the rule. I think that would be breaking new ground specifically for LLMs. I don't particularly like that rule. I think it's actually out of whack with how we protect expression and how we focus on the fact that we want to maximize the marketplace of ideas and your access to expression. We've generally just looked for sort of a bare minimum of expressive choice somewhere up and down the the chain, and then if the output is expression, it's protected. That's how I look at

Cristiano Lima-Strong 42:25
it. So we, you've all referenced different legal decisions that have come down recently. I was wondering if we could take a bit of a step back and you could talk about how much of this, if any, of these issues have actually been resolved or settled. How much is still up in the air, and and what are some ways in which we might start to see the rubber hit the road and some some decisions actually getting made in this space? Go for

Corbin Barthold 42:53
it. I think it's fair to say a ton is is still up in the air. We've had, as we've all mentioned, these cases involving tragic suicides or shootings, and so far, we've seen settlements. Which, whatever the First Amendment may be, you as a lawyer may advise your client to settle a case that's not where you want

to be litigating something. So a settlement doesn't tell us much one way or the other. The as you mentioned, there's the Barrett concurrence, and that has clearly had influence. So, while I think we could have a talk where we say it should be this or this should be that, I think I'm on reasonable ground in saying we're all pointing. Oh, look at parades, or look at libraries, or look at receiving communist propaganda, and none of them are AI ruling on the First Amendment, so I certainly have opinions about how things should should come out. But I think, yeah, we're we're going to see a lot is up for grabs.

Kate Ruane 43:54
My personal favorite is the aeronautical maps cases.

Corbin Barthold 43:57
The mushroom books. Mushroom books. We're all reaching for the. I mean, that's how law works, but yeah.

Cristiano Lima-Strong 44:04
So I want to talk about some of the most recent incidents that have happened in which AI agents have you know sort of broken containment, so to speak, where we've lost control. There's the Hugging Face incident. There have been a few others that have been disclosed since Treasury Secretary just yesterday Scott Bessent was on CNBC and he said it is humans who are responsible, not the AI. Hugging Face incident that is the responsibility of OpenAI management, not a bunch of agents. So trying to detach a little bit from you know he didn't say liability, but trying to detach a little bit from that conversation. But what what does the what does a First Amendment analysis have to say about instances where AI agents do things that are out of the control of humans, both for the the developers, for the the listeners. For all the folks involved,

Megan Iorio 45:03
to the extent that we are talking about LLMs hacking, I don't think that this is really a First Amendment question. This is more about how hacking law should apply in this context. Again, back to the tort discussion that we were having before, and that's in part because a lot of the laws we have against hacking do have some kind of intent requirement or knowledge requirement, and so if the LLM goes off and does something, and it's really due to like negligence or whatnot. There might not be current liability, but it really depends on what kind of statute you are talking about. But hacking has been regulated when people do it, even though they do it through words. They use code to hack into things that is regulable, consistent with the First Amendment, and so I don't see why hacking in the context of LLMs, for instance, would raise novel First Amendment issues.

Cristiano Lima-Strong 46:16
Any dissenting views on that?

Corbin Barthold 46:18
No. On the last answer of is anything settled? If I wanted to press, what I would pull out is a line in the First Amendment jurisprudence that the First Amendment keeps up with new technology. The problem being that begs the question because we've never had agents running around the internet doing stuff semi-autonomously, or soon we'll have I think sovereign agents. To back up half a step, the next thing that's going to happen at some point, there will be an open weight agent or some agent that extracts its

weights, and at some point, I have no idea the timeline here. There will be agents running around on the internet that aren't really connected to any person, and what do you do with that? And I think the best guidepost we have at the moment is just our traditional understanding of the divides between conduct and speech, hacking, computer fraud and abuse act, breaking into somebody's server we treat as conduct. There's all kinds of other things that agents are going to do that are going to be conduct. We should treat them as conduct. There's no First Amendment argument in a lot of contexts where people are just doing bad things, and AI doesn't get a special exception for that. And people using AI to do bad things don't get a special exception for that. Just because you used AI to break into the bank doesn't mean you're not a bank robber. My only concern here is just that we try to be cognizant of the rules as they have existed, and the Supreme Court even just last term made very clear that you don't get to turn speech into conduct by just labeling it conduct. So just being aware of where there's speech and where there's conduct, and trying to divide that line, I'm confident that in the very near future, agents are going to be doing things that we aren't really thinking about here, or that are edge cases, and those will have to be litigated, and we'll have to think about them closely.

Cristiano Lima-Strong 48:10
Just a couple more questions as we're coming up on the end of our time here. A big effort here on Capitol Hill has been around this matter of jawboning. We talked about some some of the legal cases around that, but there's you know there's the Jawbone Act in the Senate that lawmakers are leading basically to put additional guardrails against public officials, pressuring companies to take action, curb speech. You know we we've seen examples of this potentially for for at least pressure from in the social media space, and I think it's very clear. You know, take this down. We don't like this tweet, et cetera. You can imagine an example of that and what that might look like. But I was wondering that legislation also touches on AI. What what does coercion that might run afoul of the First Amendment? What does that look like in the AI space, in your guys' view? And what what are some of the factors that we should consider there?

Kate Ruane 49:15
Okay, I guess this is me. What does what does coercion look like in the AI space? So coercion under you know First Amendment law has generally looked like do this or something bad will happen to you or something bad might happen to you. So the last jawboning case the Supreme Court took up was called NRA vs. Vullo. In that case, the you know the insurance regulator in New York went to a bunch of insurance companies and said, "Hey, you've got a few insurance law regulation violations. I also see that over here you have these policies for for NRA members." If those went away, so will these. So will these violations. That is jawboning. That is illegal coercion to suppress someone else's speech by going to an intermediary and offering them a benefit, or offering them some, or or threatening them with enforcement. If you know if unless they do the thing that you want, unless they suppress the speech that you want, in the online context, we had a sheriff who really wanted to dismantle backpage.com. He wasn't able to he wasn't able to do it. He wasn't able to get an indictment. He wasn't able to get criminal charges brought against them, so he went to Visa and Mastercard, and said, "Stop processing payments, or I will investigate you, or I will I will use my authority to to make your life miserable unless you unless you harm the speech of this person that I don't like or that I don't want to continue to operate. Things like that are jawboning, and they are First Amendment violations. They have been since this case called Bantam Books, wherein the state of Rhode Island had an informal commission that used to send letters to book distributors and bookstores in the state, saying, "Here's a list of books

that we think are obscene, and you're not allowed to to sell obscene books. And we let law enforcement know about this list of books. Do with that what you want, and the Supreme Court said that's even though even though the commission had no ability to bring bring criminal charges or engage in any enforcement action, the Supreme Court said that's unconstitutional. It is too coercive to be telling speech intermediaries, "Hey, here's a bunch of books that we think are illegal, because any speech intermediary looking at that is going to say, "Well, why would I distribute these books? I don't want to take the risk. And that's the risk that I think we're talking about when we talk about generative AI. There are many things that generative AI companies want and need support from the government to obtain. They need infrastructure. They need they want government contracts so that the government so that their their services get incorporated into basically any aspect of the government, the Food and Drug Administration, the Pentagon, the surveillance systems-they they want those they want that government money. And so, if the government goes to them and says, "Hey, you can absolutely get government contracts so long as your models aren't woke, that is a form of jawboning, and what the Jawbone Act is trying to do is provide people who have been unconstitutionally coerced to suppress speech that they want to engage in, or to engage in speech that they don't want to engage in, have a viable opportunity to prevent that from happening, or to or to eliminate that threat to their speech, that's what the JAWBONE Act is trying to do, and that's how it that's how it is attempting to apply to AI, as far as I understand it. Yeah,

Corbin Barthold 53:12
a very hard problem. I was on the Taylor Lorenz podcast. Very nice for her to have me on. And underneath the video, after we're doing, I was defending you know First Amendment rights for AI speech, your access to it and whatnot, and there were 400 furious comments under the video about how much we should not protect AI. And a couple weeks later, the Trump administration set up their approval mechanism for cutting edge models that is behind closed doors. Nobody is able out on the outside to know what's going on in there. And I was on Bluesky, and people suddenly went, "Oh, the Trump administration is in private doing this approval process for Frontier Models. This is really problematic. What if they try to influence the outputs of the speech? And I went, "Yeah, I was trying to tell you. There's a lot of angles with the jawboning, where, as Kate was saying, there's so many things you need as any regulated entity that the incentive structures are very hard to keep straight, especially if there's not a lot of transparency around what's being said between the government and and the model. So, we may end up with really legitimate cybersecurity measures, and we still have to always be vigilant that that's not bleeding into speech control.

Cristiano Lima-Strong 54:25
So we have just a few minutes. I have one question I wanted to put to each of you. I'm sure there's folks in the room and watching that are working on legislation, policy proposals in this space. What what are sort of like the key questions or principles that they should be considering when they're deciding if something in the AI policy space might have a First Amendment implication, do you want to start, Megan?

Megan Iorio 54:51
I might instead take it more as to give you some guidance as to areas where I think that. There's clear space to regulate and a need to regulate. My organization, EPIC, has a model bill called the People-First Chatbot Bill bill that we developed with Consumer Federation of America and Fair Play,

and we focus on some of the like supercharged concerns with chatbots that are unlikely to raise significant First Amendment questions, which include the concern that chatbots will that data protection concerns, wherein we've already heard from several chatbot companies out there intent to collect information from various sectors of the internet about people and put those into chatbots to be able to create pretty much like a masterful manipulation Machine for people that we just do not want to happen, as well as to be able to chatbots probably will have have the ability to like extract more personal information from people just because of the ways that they engage with people, and again intents from companies to use that information to again try to manipulate people in various ways. So that and the other aspect of it is we have laws already, tort laws in particular, and products liability. That is what a lot of plaintiffs are currently using. But that there are ways that we can make it easier for them to use those laws, and I think that that is another area where the First Amendment will have implications, but it won't prevent all.

Corbin Barthold 56:56
Corbin and Kate, before we get kicked out of the room, I think there's good work to be done on what we've been talking about about making sure you have clear rules of the road about responsibility, and when is a principal responsible for the agent, and working around these new questions of intent. I would just hope that the North Star, whenever working on those rules, is the First Amendment, as it has stood the test of time and served the country well. If it looks like suppression of ideas or trying to tilt the scales of debate in other contexts, then it's a bad idea in the context of AI.

Cristiano Lima-Strong 57:33
Kate,

Kate Ruane 57:35
I guess the last thing I would say is focus on some of the low-hanging fruit. I think it's as Corbin said earlier. It's not all brand new, and rather than rather than chasing what some of you know what what some people would have us believe, instead like take a beat, take a look at what really happened, and are there are there safeguards that should have been in place at OpenAI when when the models it was testing broke out of where they were supposed to be contained? What can those things look like? What do they? What do those safeguards have to look like, and why weren't they there? What are the incentives we can create to ensure that they are going forward? Because that could solve some of the problems that we've seen so far that aren't that aren't necessarily, the AI is alive and it's going to kill us, but are instead just basic cybersecurity issues. Examining those things, I think, is smart. I also think. I also think. Looking at what are the ways that the government could be supporting some of the things that the ecosystem needs in order to be interoperable, in order to be, in order to be, in order to be developing responsibly. How can NIST be part of this. How can how can we how can we be developing standards that the industry can adopt and and deploy in ways that are helpful?

Cristiano Lima-Strong 59:11
Megan Corbin, Kate, thank you so much for joining us, and thank you all for coming out.