



## Constitutional Challenges in the Age of AI

Tuesday, May 21, 2024

Visit our media library at [constitutioncenter.org/medialibrary](https://constitutioncenter.org/medialibrary) to see a list of resources mentioned throughout this program, watch the video, and more.

**[00:00:04.4] Lana Ulrich:** Welcome to Live at the National Constitution Center, the podcast sharing live constitutional conversations and debates hosted by the Center in person and online. I'm Lana Ulrich, the Vice President of Content.

**[00:00:18.3] Lana Ulrich:** In this episode, tech policy experts join to discuss how constitutional principles apply in the age of AI. Joining the conversation are Mark Coeckelbergh, author of the new book, *Why AI Undermines Democracy and What to Do About It*. Mary Anne Franks of George Washington University Law School, and Marc Rotenberg of the Center for AI and Digital Policy. Tom Donnelly, Chief Content Officer at the National Constitution Center, moderates. Here's Tom to get the conversation started.

**[00:00:49.5] Thomas Donnelly:** Hello, everyone, and welcome to the National Constitution Center and to today's convening of America's Town Hall. My name is Tom Donnelly, and I am the Chief Content Officer at the National Constitution Center. And I'm so excited for this conversation today about AI and the future of democracy. This program is made possible through the generous support of Citizens Traveler, the nonpartisan civic engagement initiative of travelers. So thank you so much for your support. Now I get to introduce our amazing set of panelists, I really can't think of a better trio to discuss AI, policy, democracy, than the three great thinkers that we have today. The first is Mark Coeckelbergh, who is a professor of philosophy and media and technology at the University of Vienna. He's the former president of the Society for Philosophy and Technology. He's written a number of books and articles on philosophy and the ethics of technology, including AI ethics, the political philosophy of AI and his latest magnificent book, which is called *Why AI Undermines Democracy and What to Do About It*. Mary Anne Franks is the Eugene L. And Barbara A. Bernard Professor in Intellectual Property, Technology and Civil Rights Law at the George Washington University Law School.

**[00:02:01.6] Thomas Donnelly:** She's an internationally recognized expert on the intersection of civil rights, free speech, and technology. She advises several major tech platforms on privacy, free expression and safety issues. She's also the President and Legislative and Tech Policy Director of the Cyber Civil Rights Initiative, a nonprofit organization dedicated to combating

online abuse and discrimination. And on top of all of that, she's also the author of the forthcoming book due out later this year, *Fearless Speech, Breaking Free from the First Amendment*. And then finally, Marc Rotenberg is the executive director and founder of the Center for AI and Digital Policy. He co-founded the Electronic Privacy Information Center and has served on many international advisory panels. He also helped draft *The Universal Guidelines for AI*, a widely endorsed human rights framework for the regulation of artificial intelligence. Marc is also the author of several books, including the 2020 AI Policy Sourcebook and *Privacy in Society*. He also serves as an adjunct professor at Georgetown Law. So as you can see, these are three extraordinary panelists. And so I'd like to thank Mark Coeckelbergh, Mary Anne Franks, Marc Rotenberg. Thank you so much for being here at the National Constitution Center.

**[00:03:13.0] Mark Coeckelbergh:** It's a pleasure.

**[00:03:13.8] Thomas Donnelly:** So first, beginning with you Mark Coeckelbergh, and your new book, it's all about the relationship between AI and democracy. And we're gonna talk a lot about AI today. But first, I'd actually like to lead off first with a lot of what you have to say, what your conception of democracy is, to table set for us, and sort of some of the big principles that underlie that concept.

**[00:03:36.0] Mark Coeckelbergh:** Yeah, that's a really good question because it runs through the book. So my view is that democracy is more than majority voting, right? So many people think of the representative systems that we have and then majority voting within that. And so for me, I went back to ancient philosophy and try to retrieve this more Republican, philosophical Republican idea of freedom, according to which of democracy, according to which democracy is also about participation of citizens and not just voting. Participation, deliberation, and that goes then towards the end of the book, towards some kind of conception of democracy as deliberation about the common good. So that's a much sort of richer ideal of democracy, but I think one that we need to cope with the current challenges in the digital age.

**[00:04:41.7] Thomas Donnelly:** Excellent. Thank you so much for that, Mark. Marc Rotenberg, we've placed a little bit about democracy on the table right now, I'd like to turn to you to maybe talk a little bit about just some level setting on AI. And you as one of the world's experts on AI and policy, it'd be helpful if you could give us, our audience, maybe a quick state of AI address, sort of where we are right now with AI and what's down the road in three months, six months, et cetera.

**[00:05:11.4] Marc Rotenberg:** Well, it's very difficult to predict, it is a rapidly evolving technology and I think the one area of agreement is that it is having significant impact on industry, on government, on our society. I would like to call attention to the primary report that my organization produces, it follows nicely from Mark's earlier remarks, and that concerns the

alignment of AI policies and practices with democratic values. Because we are specifically interested in trying to assess what direction this technology is taking us, as reflected in the policies that national governments adopt. And to Mark's point about public participation in a democracy, when we set out our methodology of 12 metrics defining democratic values, we actually considered process-related issues, such as does the public have the opportunity to meaningfully participate in the development of national AI strategies? Are documents relevant to government deliberations readily available to the public? So we think about democracy not only in terms of the protection of rights and the structure of governance, but also for the ability of the public to participate. And I would say that as we look at this constantly evolving technology, it's actually very important to be anchored in these factors and these considerations.

**[00:06:44.5] Marc Rotenberg:** Because it's too easy to get swept along with what the technology may make possible. I would suggest without considering what we consider to be valuable.

**[00:06:55.8] Thomas Donnelly:** Excellent. Thank you so much for that, Marc Rotenberg. And turning now to Mary Anne Franks, you're one of the nation's leading experts on the First Amendment, on things like online speech. And so to sort of set the table for our discussion in that realm, can you just give us an overview of some of the big constitutional debates we're having right now over free speech, online, First Amendment, online speech and sort of where AI intersects with some of that?

**[00:07:26.3] Mary Anne Franks:** Absolutely. And when we think about the First Amendment, and we think about it a lot, I think in this country, when we are debating what values really matter to us as a democracy, and freedom of expression is obviously central to that. But very often we have this tendency to think of freedom of expression as being sort of completely covered by the doctrine of the First Amendment, when in fact the First Amendment is quite limited in its scope, because it's just about what the government is not allowed to do vis-a-vis the rights of citizens, the rights of the people to speak. And so I think what we're seeing with this new evolving landscape and AI and other challenges from technology, is kind of interrogation and a kind of highlighting of all of those limitations that we've seen the doctrine really have. Including what does it really mean to speak in the digital age? Who counts as a speaker? Does it have to be a human source that is behind any particular form of expression? But even more than that, those underlying values that we care about when it comes to the First Amendment and free speech as essential for democracy are usually thought of as being autonomy, truth and democracy, right? That accountability and transparency.

**[00:08:39.5] Mary Anne Franks:** And when we think about the ability of deepfakes, digital technologies that make people think that something that is false is true, or equally important to make people think that something that is true is false, it's really pushing up against those values.

You can't really have individual autonomy in a population that doesn't know what reality is, if we're not sure what the truth actually is, you can't really have a marketplace of ideas and any kind of access to truth, if the marketplace is constantly flooded with fraudulent ideas that are almost impossible to separate from true ones. And all of those things are going to lead into the question of civic participation and holding our political officials accountable. The possibility of new technologies to deceive us and for the population to want to be deceived is really pushing on every value that we might care about when it comes to freedom of expression and it really is creating a moment, a long overdue moment probably. For us to think about the limitations of the First Amendment when it comes to being a guideline for what we value in terms of freedom of expression and what are the best ways forward in terms of trying to preserve those values. Because some of those attempts to preserve those values might look like things that actually go against what we would superficially understand the First Amendment to protect.

**[00:10:01.0] Mary Anne Franks:** So sometimes we're going to have to look at regulating more speech in order to make sure that we can actually respect those values that freedom of speech and the First Amendment are meant to safeguard.

**[00:10:11.5] Thomas Donnelly:** Excellent. Thank you so much for that framework there. Mary Anne it's great to think about here. Mark Coeckelbergh, returning to you and to your book, one of the concepts that you have in your book dovetails so nicely with what Marc Rotenberg and Mary Anne Franks have already said, and that's this idea of creating a democratic AI. And so here you say you argue that AI needs to be rendered less damaging to democracy and preferably should work for democracy. Can you talk a little bit more about that vision and what that might look like practically on the ground?

**[00:10:45.0] Mark Coeckelbergh:** Yeah, the vision is not only that philosophical republicanism, but also the ideals of the French and American revolutions where you have, of course, the value and principle of freedom which is important, but also values like justice and fraternity. And I think these are also fundamental values for democracies. Now, how then to do that with AI, creating democratic AI, is a big question. But, so one way is, as Marc suggested that to involve citizens in deliberating about the future of AI, which is really needed in a time when societies where big tech really decides our technological future. I think we need to have democratic involvement in what that future is gonna look like. But for me, it's also about education, as Mary Anne said, we need people who actually know some things, I think for me, this is really important in a democracy to get that to work. And then in terms of practical arrangements, I think there are two things at least for democratic AI that are needed. One is that there is some kind of involvement of these democratic values that they are somehow inserted in the development process. So at the technical level, this is where things happen and where decisions are made at the level of companies and their development of the technology.

**[00:12:24.8] Mark Coeckelbergh:** So we need to really make sure that we insert those values there already, because otherwise we have problems afterwards and then regulation it's not always that effective to deal with the huge challenges. Another concrete proposal I have is to do something about the balance between expertise and democratic legitimacy. Currently, there's a sort of danger of some kind of technocratic rule by these big companies, and I think we need a sort of balance between expertise and the democratic legitimacy. And especially not only a balance, but some kind of meaningful link between the two. What's happening now is that ad hoc, we have this new technology and suddenly we need, we have hearings, for example, in parliaments, in Congress, and people like Sam Altman turn up there. But what we really need, I think, for the future is some kind of more continuous way and meaningful way of connecting experts with political representatives and with the democratic institutions. And that's a challenge itself, too, because it means institutional change. But I think we cannot just pretend that we in the 21st century can just have those 19th century parliaments and think that this is enough. We live in an age where really expertise is needed and that we need to do it in a democratic way.

**[00:14:10.5] Thomas Donnelly:** Excellent. Thanks so much for that, Mark Coeckelbergh. There's so much to drill down on in that answer that we'll return to hopefully as this conversation proceeds. But returning to you, Marc Rotenberg, I really wanna do as much as possible in this conversation to sort of get down to the ground level and figure out what is happening right now with AI policy. Sort of what are some of the big concerns that are animating policymakers. But as you've noted, just looking in the United States, AI is one of those areas that despite our divided politics, we see a lot of bipartisan bills out there in Congress. So maybe if you can give our audience a sense of what sort of concerns are maybe helping spur that bipartisanship, what are sort of some of the common concerns, and also what are some of the the policies that are being concretely considered inside Congress today.

**[00:14:57.5] Marc Rotenberg:** Sure. Well, the list is long, actually, we can go from algorithmic transparency, to misinformation, to election integrity, job displacement, climate change, interestingly, because we're now aware that these large language models require a great deal of energy. If you had asked people, let's say, 10 years ago, what are the cornerstones for AI policy, people would have said things like fairness, accuracy and transparency, right? As key goals. But as more applications have developed and as we see the industry evolving, we now look at issues including equity and gender equity. What are the representation and data sets as these automated processes are making recommendations? I also wanted to briefly come back to your original question to say that, I've been in the AI field for many, many years, I'm afraid to say quite how long. But before I went to law school, I was writing computer programs for chess and backgammon, and we actually called that a form of artificial intelligence. But these were simple rule-based expert systems, and in fact, they were being developed for engineering and medical applications, and they were readily understood, and most critically they were provable as to the outcomes.

**[00:16:25.3] Marc Rotenberg:** The results were replicable, we could trace an outcome back to the coefficients and say, Oh yes, that's why the chess program recommends moving the knight to this square and not a different square. Today, we're in a very different world with AI, we've moved from these rule-based deterministic systems to probabilistic systems, stochastic systems that produce recommendations based on an enormous amount of data, but you look at an output and you actually can't prove it. And this is cross-cutting with regard to policy, it has impacts in housing decisions, employment, credit, criminal justice, all across the board. So I would say in this last category, as we're looking at the policy responses, we're looking at efforts to establish forms of accountability for what has basically become opaque decision-making. And here I would say the European Union, perhaps not surprisingly, both with the GDPR and with the new EU AI Act, has now established legal standards to ensure not simply explainability, which can be a rationale for an outcome, but actually access to the logic data and factors that produce a result that impact people. And this is, in fact, I would say a very important development.

**[00:17:56.8] Thomas Donnelly:** Excellent. Thank you so much for that, Marc Rotenberg. Returning to you, Mary Anne Franks, you've done a lot of writing about the effects of AI on, in particular on women and girls. You've said that women are canaries in the coal mine when it comes to the abuse of artificial intelligence. Can you talk a little bit about that and sort of why you've placed focus there?

**[00:18:21.0] Mary Anne Franks:** Yes. When we first started using, or first started hearing the term deepfake, which is kind of a casual term that people use often to refer to manipulated imagery that is so virtually indistinguishable from a real image that people are fooled by it. That term comes from a user on a message board that was essentially democratizing the technology and making it available to the public to create these highly realistic pornographic videos and images of celebrities is where it started. And then it's now grown, I would say to target basically anyone, so it's female celebrities primarily targeted. And now there's been this explosion in the last few years of applications, websites that are just about generating pornographic depictions of women and girls without their consent, famous or not famous. And we've seen it obviously all the way from really super famous people like Taylor Swift to high school girls in New Jersey. And I think one of the things we have to think about when we observe a phenomenon like that is, why are we in this particular state now? And what I mean by that is, was there any way that we could have been better prepared for the unleashing of this kind of technology for these malicious purposes?

**[00:19:40.2] Mary Anne Franks:** And my suggestion is that there were earlier versions of what we now call deepfakes, cruder versions, more superficial versions that had happened in the 1990s when the internet starts supporting visual images. One of the first thing people use it for is to do these really crude photoshopped edits of female celebrities onto pornographic depictions. And

those celebrities, those individuals who were targeted were unhappy about this, were trying to say this is a harm, this shouldn't be allowed. And it was dismissed as being just not that serious of a problem, something that is just what female celebrities have to take as a price of their fame. And what I want to point out about women and girls being canaries in the coal mine is that when abusive technologies develop, they're usually tested on women and girls first.

**[00:20:27.9] Mary Anne Franks:** And I think that development we've seen of this kind of imagery that has no respect for consent, no respect for context, has no concern for the kind of harms that can be imposed when you have these depictions. That's what's led us to this moment. And if we had taken seriously that kind of image exploitation in its cruder forms in the '90s, we might have had better legal structures in place, better guidelines in place for the technology that we're now facing, and which now people worry about because it can also be used for political purposes. We can now have deepfakes of politicians taking bribes when they didn't. Or we could have deepfakes of politicians who are not taking bribes when they did. And so now we're seeing in real time, you know, just how serious and wide ranging the consequences are. But we had plenty of notice.

**[00:21:18.7] Mary Anne Franks:** It was just that the people who were primarily being harmed were women and girls. And the way that they were being harmed was a form of sexual exploitation. And like in so many other instances of the law, the law just didn't take it that seriously. So I think there's a lesson here, in addition to an underscoring of the urgency of who today is being most affected by malicious uses of AI, it's primarily women and girls and other vulnerable communities. And it's also a lesson for saying whatever we're going to have to face as an entire society and the foundations of our democracy is almost always going to be weaponized against those vulnerable groups first. So we need to take that seriously when they happen, both because they deserve that respect when they happen to those communities, but also because those weapons are going to be turned on all of us eventually.

**[00:22:06.1] Thomas Donnelly:** Excellent. Thank you so much for that, Mary Anne Franks, for that urgent warning. And that's an excellent thing to continue to keep in mind. In terms of you, Mark Coeckelbergh, what part of your book that is notable is that you really explore the history of what you describe as the problematic relationship between technology and democracy. So on the one hand, you try to provide sort of a theoretical framework for thinking about debates over AI and democracy today, but you also bring in sort of an historical narrative. What can we learn from that history? And what light does it cast on some of the challenges that we face today when it comes to AI and the future of democracy?

**[00:22:44.5] Mark Coeckelbergh:** Yeah, one thing we can learn is that already in ancient philosophical thinking in Plato, there is the idea that you need expertise to steer the state. And he uses this metaphor, which happens to be basically cybernetics, right? So he already proposed a

political cybernetics where expertise is key. And what's right about that is that you obviously need some expertise in politics, and that's for sure true in the kind of knowledge societies that we have today and in these highly technological societies. On the other hand, his idea was already dangerous back then, and dangerous in a sense of anti-democratic. So here cybernetics was already linked to an anti-democratic idea, And then if we look also in the history from ancient civilizations to today, you see that technology is often used to support the most powerful in society and to centralize power.

**[00:23:54.1] Mark Coeckelbergh:** And that happened, let's say, in Egypt already, but it also happens today when these big companies continue to accumulate more power. So there's a history of technology and centralization, and AI kind of fits very well in that history. I think it gives us a warning that we have to watch out that we don't use and develop AI in a way that contributes to further centralization of power and anti-democratic tendencies. And I think, you know, the example that Mary Anne, an example that Mary Anne just gave are very much linked to that. You can't have a democratic society when you exclude certain groups. You can't have a democratic society when power is totally asymmetrical and is held by small elites. So these are, I think, good lessons from history which we should keep in mind today in AI policy and AI governance.

**[00:25:01.0] Thomas Donnelly:** Thank you for that, Mark Coeckelbergh. Returning to you, Marc Rotenberg, I'd like to bring in one of our audience questions here, which sort of follows on with your previous response. It's from Niels Nielsen, and here's the question. Governments and businesses are using machine learning to qualify people for loans and grants, screen job candidates, detect fraud, and make other critical decisions. The opaque nature of these machine learning models means we may never know why we were denied a benefit or even charged with crimes. What changes in the law are needed to assure that there is always an effective human override to computer-based decisions?

**[00:25:35.4] Marc Rotenberg:** Well, it's a very good question, and I think the question also suggests the answer, that we do, in fact, need changes in law to establish transparency as to decision-making and to maintain human control over the outcomes. It was actually almost a decade ago, as these issues were coming forward, I launched a campaign in support of algorithmic transparency. And we were basically arguing in the moment that as the decisions become deeply embedded in statistical systems, it's not only the person who's impacted by the decision, but actually the organization that's responsible for the decision that may not see the basis of the outcome. This happens, for example, in the employment sector when large companies delegate to third-party vendors the initial screening of resumes and the vendor comes back and says, of the 10,000, we recommend you talk to the 200.



**[00:26:35.5] Marc Rotenberg:** The criteria may not be known to the company that has selected that vendor why precisely those 200 were selected. Now, as I said, if you look to the EU, you see in the GDPR already in Article 22, a requirement that for any automated decision-making that affects a fundamental right or has a legal effect, a person should have the right to a human determination. This is not the ads that you see when you browse the internet and it's not how your video games operate. But if you are denied a job under the GDPR, you could be entitled to a human review.

**[00:27:18.8] Marc Rotenberg:** And this is actually carried forward also in the EU AI Act in Article 8, which speaks of access to the logic and the reasoning. And I think these are positive developments. I'd like to see them adopted in the US and elsewhere. I think ultimately, you know, fairness is in the interest both of the person who is subject to a decision, but it's also in the interest of the decision maker. We should make decisions that we can stand behind. And not turn the decision-making process over to a system we don't understand.

**[00:27:53.8] Thomas Donnelly:** Excellent. Thank you for that, Marc Rotenberg. Returning to you, Mary Anne Franks, you spoke so powerfully about the threats and the harms created by deepfakes. Are there certain policy solutions on the table to address deepfakes? Do you think they're likely to work? And then, you know, the last part being, are there any meaningful constraints on what sort of government action we can have in that context in light of First Amendment doctrine, etc?

**[00:28:23.0] Mary Anne Franks:** Well, yeah, that's the complicated question, right? There are some really interesting proposals currently. There is a really good bill that is sponsored by Representative Morelli of New York, and that's in the House, which would directly focus on the question of sexually explicit digital forgeries and prohibit them through the criminal provision and also a civil provision that would provide a remedy to victims. So that's really important. And that's the kind of legislation I think we really need to see. And I particularly like Morelli's bill. And I was involved in advising his office because it has such a careful definition of digital forgery, a careful definition of what's being targeted because of the First Amendment issues that we have to keep in mind.

**[00:29:10.0] Mary Anne Franks:** We don't want to impose a regulation that is going to inhibit protected speech, but we also want to recognize that there are certain forms of false speech that cause very devastating, irreversible harm, that we have a history in this country of being able to say that kind of conduct, that kind of speech can be regulated and should be regulated. We have fraud laws, we have laws against counterfeiting, laws against identity theft, impersonation. So very much in that vein, I think, is what we need to see here for image exploitation through the use of AI. But there are constraints, right? Because not only do we have to be very careful and narrow about those, we also have to ensure that Congress is actually going to pass any of these.

**[00:29:53.8] Mary Anne Franks:** And as was mentioned before, there is, I think, some cross-political agreement about how AI is posing some problems that we should address, but the devil is always in the details and there may be very difficult political headwinds to meet. We've been trying to pass legislation on the disclosure of authentic sexually explicit imagery for going on almost a decade now, and Congress has not yet passed it. So we clearly got these obstacles still in place from people who I think sometimes have this knee-jerk reaction that any attempt to try to say there is confusing, fraudulent, harmful conduct that should be regulated, their first response is to say, well, you can't because of the First Amendment, which isn't true, but has a lot of power. So that's one of the difficulties there.

**[00:30:39.4] Mary Anne Franks:** And I think the other major difficulty is the incentive structure for the tech industry. And here, invoking what I know the people on this panel know well, but the particular federal law in the United States called Section 230 of the Communications Decency Act, which has been really broadly interpreted by courts to effectively insulate them against essentially anything that happens on their platforms, with some exceptions. I think if we're going to actually expect that the major players here, because even if they're not the ones producing this kind of abusive imagery, they're often the distribution points, right? So if deepfakes get made of a celebrity via a certain website or an app, they get distributed on a place like X or Twitter or what have you. And if there's no incentive for these companies to remove that in a timely manner because they've got protection from suit, then we're gonna see them not take it down particularly quickly.

**[00:31:30.9] Mary Anne Franks:** So there's this larger problem that I think the AI issues are underscoring, which is the incentives for the tech industry are backwards. We don't have nearly enough accountability for the industry. There's far too much encouragement for the industry to simply be as reckless as it wants in its pursuit of profit. Because engagement is everything and engagement is the euphemism for what is often extremely polarizing, harmful, extremist, exploitative content that people are going to click on. So. We're going to certainly have to have carefully crafted legislation directed at certain abuses that comport with the First Amendment, but also we're going to have to do something radical, I think, to change the incentives of the tech industry to try to design platforms and services that actually value truth and let it rise to the surface, as opposed to flooding the information ecosystem with fraudulent material.

**[00:32:22.6] Thomas Donnelly:** Excellent. Thank you so much for that, Mary Anne Franks. Turning to you, Mark Coeckelbergh, you know, one of the things you talk about in your book is how AI can present, you know, serious threats to democracy, especially when democracies are weak. And so here's what you write. You know, when our democracies are already weak, the negative effects of digital technologies such as AI will hit harder. Weak democracies are less resilient against what AI can do to them, and AI amplifies the problems and in this way supports

anti-democratic tendencies that could potentially lead to authoritarianism. Can you just please talk a little bit about that observation and some of those challenges that AI presents in this context?

**[00:33:02.7] Mark Coeckelbergh:** Yes, I think if you already have a system where, for example, populist tendencies are very, very easy to get off the ground, where there is not much education or somehow the education is very unequally distributed. When there is a rate of discrimination, when there are parliaments that do not really function effectively, then I think AI can be used as an instrument easily to further undermine democracy. And in the book, I discuss specific ways that AI undermines particular principles, including freedom, but also principles like justice, when there's discrimination, or also talk about these asymmetries in terms of power. In that sense, I think one of the things that needs to be done is to, in the background, next to passing laws, which I'm sure that many people are waiting for in the US, next to that, I think in the background, democratic systems also in US need to be strengthened.

**[00:34:25.1] Mark Coeckelbergh:** I also sort of paraphrase Gandhi in the book and saying like just as Gandhi said that civilization would be a good idea I say like you know democracy would be would be a good idea in the sense that you know we don't have full democracy yet. So I think we need to strengthen those institutions and also strengthen them against the power of big tech if we don't do that we create sort of background conditions for less democratic situations and for sure less democratic use of AI.

**[00:35:06.0] Thomas Donnelly:** Excellent. Thanks. Thanks, Mark Coeckelbergh. Marc Rotenberg, bringing in another audience question. The audience is hungry for information about AI policy here. We have a broad question about, are there any current laws being considered by Congress to help rein in misinformation from AI?

**[00:35:25.1] Marc Rotenberg:** That's a good question. I think there are actually several bills on misinformation, particularly in the election context. It was, in fact, this morning that I followed the hearing in the Senate Rules Committee where Senator Klobuchar has introduced three bills regarding election integrity. They're bipartisan, by the way, so she has Republican sponsors on all of them. And the aim is to address concerns about the use of deepfakes in elections and other types of representations that could undermine the integrity of an electoral outcome. I will say there was a bit of sparring in the committee because in the US of course, a lot of the administration of elections is left in the hands of state officials. But with regard to federal elections, the federal government certainly has the legal authority to establish necessary standards.

**[00:36:25.8] Marc Rotenberg:** One of the ways in which misinformation could be addressed or AI-generated information is through watermarking techniques. Now, in talking about

watermarking, I need to say there's also a fair amount of skepticism because oftentimes when a technological solution is proposed to a technical problem, someone almost immediately comes up with a way to defeat the solution, which isn't to say the solution still doesn't help if you can identify the majority of AI-originated content through watermarking. And in fact, in the AI executive order that President Biden issued last year, there is clearly a push toward watermarking. But I guess my final point on this issue, and looking, of course, to Mary Anne, who's an expert on all of this. We do need to be careful about the term misinformation.

**[00:37:21.1] Marc Rotenberg:** It can be highly subjective, and as a general matter, the First Amendment doesn't allow the prohibition on speech simply because someone says it's incorrect. But at the same time, I think we can also state that as to who the speaker is, we should be able to distinguish between speech that originates from a natural person or legal person and speech that is generated by an AI. There are many who argue that AI is simply an extension of what a person chooses to produce, but I think that's becoming increasingly less clear. And what we hear more and more from the world of AI science is that we're looking at systems that are becoming increasingly autonomous in their operation. Whether they are given legal autonomy or legal agency is an entirely separate matter but how they operate and the outputs they produce, I think, is going to make this question of misinformation much more complex.

**[00:38:23.6] Thomas Donnelly:** Thank you, Mark. And Mary Anne Franks, given your expertise, please do feel free to follow up on anything that Marc Rotenberg said in his response there. And more broadly, as we think about deepfakes and about the threats of AI in other contexts, maybe talk a little bit about what you envision being the relative role policy-wise of the national government versus what role the states might play, versus even what role individual-level litigation under existing law may play in addressing deepfakes or, you know, sort of any of the other big threats you see AI threatening, threats exist with AI and democracy more broadly.

**[00:39:08.0] Mary Anne Franks:** Right. So I do think to underscore those points about how one of the reasons why we need to be concerned about manipulated imagery and audio, all of these kinds of issues that are kind of housed under this term of Deepfakes is because as Hannah Arendt had pointed out many years ago, when you have a society that cannot tell truth from falsity, you can do anything with them. And the dangers there for authoritarianism and for the undermining of democracy, or the continued inability of democracy to fully assert itself maybe is a better way to put it. These are really serious issues that require careful attention. The limitations, I think of some of the legislative approaches that it's certainly true as Marc Rotenberg was saying, that putting emphasis on watermarking and things of that nature will I think, mitigate their partial responses to some of the problems that we're facing.

**[00:40:01.0] Mary Anne Franks:** But as he also pointed out, there's limitations to this. Not only because the technology always races ahead of anything that the solution will be, but also because

of the way that Deepfakes operate on our cognitive and emotional the reactions that we have to it. So, two things I'd wanna point out about that. One being, when we're talking about sexually explicit digital forgeries, the fact that something's watermarked or indicated clearly that this is a fake, you go to a website that says all these things are fakes, does not actually undo the harm, right? It may help clarify that this is not a real image. This is a manipulated image, but the exploitative aspect of this, the dignitary harms, all of those kinds of harms are still there even if you openly disclose this is a fake. The second thing I would say is that even beyond the sexually explicit forgery question, Deepfakes generally, so depictions of maybe a peaceful protest as though it were violent, or again, a politician doing something or saying something that they never did.

**[00:41:02.9] Mary Anne Franks:** The question isn't just about whether you can go back and correct this. And again, AI isn't new here. Misinformation has been around forever, right? It's much more sophisticated now. But there was always a challenge to try to correct misinformation, not only because sometimes we disagree about what misinformation is as opposed to strident disagreement or editorialization or opinions, but also because as a cognitive matter, when we try to tell people this thing is false, what often ends up happening instead is that we reinforce the false message. So we know from behavioral science that there's this thing called the illusory truth effect, which means that once we see something, especially if it's highly realistic, even when we're later told that that thing is false, if someone puts that image back up and says, this is a false image, what seems to happen is that our brains basically process that as a repetition of that same image in a kind of truthful sense.

**[00:42:00.9] Mary Anne Franks:** So we don't retain the correction. We don't retain, this is false, we retain the image. And so I think that's important to keep in mind in terms of limitations of technological solutions, watermarking, provenance questions, all of which are really important to solve pieces of this problem cannot across the board solve this because of that illusory truth effect. And I think all of them are also just sort of moving around the bigger question, the bigger democracy question, which is how do you make an effective misinformation diagnosis of misinformation system when first of all, there's all these technical challenges, but there's also the challenge of how do you make the public want to know the truth? One of the problems I think that we're facing is not just that people have a hard time distinguishing between what is true and what is false, we're also dealing with the fact that lots of the population don't want to know whether something is true or false.

**[00:42:57.7] Mary Anne Franks:** They want to know whether something validates their particular worldview. So they're eager to indulge in something that might have questionable provenance because it supports something that they think is right, and they're happy to share that and to click on that and engage with that because it supports their worldview. So there's this question of why do we see so much of the problems we're having right now when we're trying to

figure out the solutions to this problem? We also have to try to understand the psychological, behavioral, political reasons why so many people are investing in falsity and want to be invested in falsity. And how is it that we encourage a public to actually want to know what is true?

**[00:43:31.5] Thomas Donnelly:** Absolutely. Mary Anne Franks, and that dovetails so nicely to a particularly ambitious part of Mark Coeckelbergh's book, which is, Mark, you argue that to address the problems of AI and democracy we have to go beyond just institutional reform, government regulation, technological innovations, that in part what we really have to do is transform culture. And you've called for whether it's a new renaissance, new Enlightenment I'm curious if you could fill out that ambitious vision for us of what that sort of cultural transformation might look like and also how we might bring it about.

**[00:44:12.4] Mark Coeckelbergh:** This is a very good question how we might bring it about. First about vision, I think that if we look at the history, we see that in the renaissance new technologies, in particular, the printing press new communication technology what was used to launch cultural project to the Renaissance. And that brought together scholars, but also in the end until today has been democratized. And there has been sort of, literacy has been really transforming also the entire society. So if we could do something with digital technologies that's similar where digital technologies are really linked to a wider societal and cultural project where literacy of people is stimulated and where people are trained in critical thinking in also developing critical relation with the new technologies, I think then we're sort of moving more towards democracy because we're creating the kind of citizens that we need for democracy.

**[00:45:32.7] Mark Coeckelbergh:** So if we can do that, for example, but I mean, there are lots of people working on digital literacy and media literacy. There are people thinking about how to reform education. If we could combine efforts towards democracy and democratic AI with those efforts and do something politically also about education, I think we are on the right track because we are creating those background conditions that people indeed care about the truth and avoid also the things that Mary Anne just mentioned in the book, I also mentioned Aren. I think it's very important to create those background conditions where truth is seen as important where critical thinking is a normal thing to do. And then it's possible to discuss with one another. We don't have to share everything. We don't have to be the same. We can keep our differences, but we can also come together and try to reach some shared understanding. And I think that's also in line with what the enlightenment both in Europe and in the US want and so in that sense, a kind of constitutional project.

**[00:46:53.7] Thomas Donnelly:** Excellent. Returning to you, Marc Rotenberg we've talked a bit about Congress we've talked a bit about the EU. If we think more broadly, I mean, what role do you see international organizations say the United Nations playing in sort of trying to organize sort of a global approach to AI?

**[00:47:15.6] Marc Rotenberg:** It's a very good question, Tom, actually. I've spent the last several years working with many international organizations on the development of global norms for the governance of AI. We began with the OECD, issued the OECD AI principles in 2019, which is first global framework. The UNESCO recommendation on AI ethics came out in 2021. We mentioned the EU AI Act. The Council of Europe is about to finalize, in fact, this week, the first AI treaty, which is quite important because the mandate of the Council of Europe is to uphold fundamental rights, democratic values, and the rule of law. So these are rapidly emerging governance frameworks. You mentioned the United Nations. There are many, many projects underway there. We're expecting a report from a high level advisory board on internet governance in a few months. I've put forward my own proposal in a recently published article.

**[00:48:18.0] Marc Rotenberg:** I've recommended that the United Nations establish a Special Rapporteur on AI and human rights. The Special Rapporteur mechanism is one of the ways that the UN has found sort of an agile response to some of these emerging challenges, which I think would be particularly appropriate in the AI world. So the answer to your question is there's actually quite a lot underway at the global level among international organizations, and almost immediately we start asking questions about harmonization, convergence, divergence what will the outcome be when companies are subject to multiple governance norms. My hope is that there will be upward direction in the regulatory structure as governments more fully understand the risks. They are more likely to legislate, they're more likely to implement the US is I would say not in first place at the moment. Maybe that's an understatement. But we've made a lot of progress in the last year, and that's encouraging. When the Senate issued today, the AI issues report, I saw a lot of good recommendations in there. I was a bit disappointed that there was no reference to many of the international governance frameworks, including frameworks such as the OECD and UNESCO that the United States has supported.

**[00:49:46.5] Thomas Donnelly:** Excellent. Thank you for that. Marc. Returning to you, Mary Anne Franks we've talked a lot about policy to this point. I'd like to return as we get to the end of our conversation here, to sort of broader questions about First Amendment theory and doctrine, which you teed up at the beginning of our conversation has come through in a number of your answers today. And just thinking broadly as we're thinking about this swartz of policy approaches you think are necessary to address things like Deepfakes, potentially misinformation online or other challenges ahead when it comes to AI policy what parts of the traditional approach to the First Amendment and First Amendment doctrine do you think sort of are well matched to that universe of concerns and what parts of it do you think we really have to rethink in fundamental ways?

**[00:50:38.1] Mary Anne Franks:** Yeah, and that's really such a complicated, and I think the really interesting question that we're facing right now that I think I've referenced before that this

is an interesting opportunity to think about the limitations of the First Amendment. And when I say that, I mean in the sense that we have to recognize that there's a kind of, not just that the doctrine may be narrow in ways that are a product of the fact that the First Amendment is quite old. And that we have seen a lot of changes in modes of expression and questions about what should constitute incitement or harm. And maybe the First Amendment hasn't really caught up as a doctrine, but even in all of its complexity, there's this kind of disconnect between doctrinal narrowness and the public's perception of what the First Amendment does.

**[00:51:22.0] Mary Anne Franks:** So it's interesting in a conversation about misinformation to have to underscore that there's so much misinformation about what the First Amendment does that is also complicating this debate in a similar way that you kind of pseudoscience that is complicating scientific discussions. Pseudo-law really interferes here because many people think that the First Amendment means that there can't be any kinds of restrictions on anything that could be conceptualized as speech, and that's simply not true. If you look at the historical categories that the Supreme Court sometimes speaks of, there are pretty big categories. Obscenity is a huge category. Fraud is a category. Defamation is getting partial protection, but not complete protection. Child sexual abuse material is an exception. And then we look at all the things that aren't highlighted as exceptions, but are things like privacy regulations, perjury laws, the kinds of restrictions that we have on FDA labeling.

**[00:52:15.2] Mary Anne Franks:** There's all these ways in which what we're really saying sensibly in actuality with our First Amendment doctrine is you can regulate and you should regulate speech when we're talking about the balance of harms. If speech is going to actually cause concrete verifiable harm, and the benefit is actually quite low, that's really the only way to explain why we have the categories of exceptions that we've had, and we've always had them. There's never been some kind of absolutist true free for all in this country about all speech all the time. And so the only question should be, what is the most responsible, socially beneficial perception, conception of freedom of expression with these new forms of technology that preserves those underlying values, again, of allowing people to have autonomy, allowing people to have a search for the truth that is possible, and allowing people to be able to hold their representatives accountable and learn what political officials are doing through democratic accountability, and ensuring that what is happening in technology or elsewhere isn't subverting all of those principles and making all of them impossible to uphold.

**[00:53:22.8] Mary Anne Franks:** So I think really just being careful about not falling for what I think is often the perceived pseudo-law of the First Amendment, and seeing what it really has done in practice, which is to intervene in these spaces where there is massive social harm and collective harm. That's something where we really need to be open minded about that and expand that notion of what counts as harm. Because linking this with some of the other comments I've made before, part of the deficiency of current First Amendment doctrine and law



generally, is that we haven't always taken everyone in a democracy. We haven't always taken their rights and their obligations equally seriously. So part of the correction we have to do in our First Amendment doctrine and elsewhere is to consider that harms that disproportionately affect women and girls, for instance, or disproportionately affect racial minorities or the disabled or sexual minorities, that we have to start taking that seriously.

**[00:54:19.7] Mary Anne Franks:** We have a long history of exclusion and exploitation of certain vulnerable groups, and that means that their concerns and their ability to take advantage of certain freedoms that other people get to take for granted that they have been compromised. So I think what we need to think about is both, what is the First Amendment done for the most powerful people in society? How can we ensure that those same protections are given to the really vulnerable members of society, even if that means in some cases, that we have to take the harms that are inflicted upon vulnerable populations more seriously than whatever marginal benefit powerful people might be getting from those forms of expression. And that's a really hard and complicated conversation to have, but it's one that is desperately overdue, I would say.

**[00:55:02.1] Thomas Donnelly:** Absolutely. And we're coming towards the end of our conversation here today. Maybe if I could get each of you go around the horn one more time if we're having this conversation again, say in 10 years and you wanted to sort of summarize an optimistic vision of where we might be in terms of whether it's policy, culture, law, whatever you'd like to focus on I'd love to end on an optimistic note if we can and maybe beginning with you, Mark Coeckelbergh.

**[00:55:32.5] Mark Coeckelbergh:** Yeah, I mean, If I can give an optimistic view, then I would say that in 10 years I would like to see that all over the world, that countries pass legislation that give some minimal protections to citizens in the light of the possibilities of AI and that protects also our democracies. And part of that legislation, I think should be some reforms to education and regulation of media in such a way that yeah, citizens are more empowered for contributing also to democracy. So I see this as possible and yeah, if you see for example, in Europe that there is already, like legislation on this, I think it's possible somewhere else too, maybe in the American Congress could be something like an AI Act light.

**[00:56:41.7] Mark Coeckelbergh:** But in any case I think that it's possible. So we can be optimistic about that. Also seeing all the efforts that are currently being done. I also hope and that's the last point, that global governance of AI comes off the ground. I think as Marc said, there are already many efforts in international organizations towards that end. Currently that's still soft law. What I would hope for optimistically if I have to, then I would hope for some minimal norms that we can agree on some kind of agency that can monitor also developments in AI worldwide. Because I think we do need as I've argued in an article, we do need global governance of AI for moral and political reasons, moral reasons to protect citizens in various

ways and also politically to make sure that AI remains accountable and that those power issues are addressed.

**[00:57:48.4] Thomas Donnelly:** Excellent. Thanks, Mark Coeckelbergh, same question to you, Marc Rotenberg an optimistic vision, if we can.

**[00:57:56.4] Marc Rotenberg:** Okay. Well, let's begin by agreeing that AI is a human creation. And as Thomas Edison said, what we create with our hand, we should control with our head. I think we need to engage this debate about AI governance with both purpose and ambition recognizing that this is a transformative technology and to ensure the benefits that many people are seeking. We can't just sit on the side and see what happens. I think we made that mistake actually with the internet economy in Section 230, I don't see any support for a Section 230 approach to AI-based services. But I also think we need to think critically about how we understand the challenge. Many people say, for example, that there should always be a human in the loop. I think this is fundamentally a misunderstanding of what the challenge is.

**[00:59:00.9] Marc Rotenberg:** If you sit with experts, with judges surgeons and others, they will say, if I'm in the loop with an expert decision system in my field, there's not very much I'm going to be able to say because we will defer to the machine. I think we need to keep humans in charge and allow AI in the loop if it has been adequately tested, assessed, reviewed, and proven, like with all devices and with all services. It's ultimately a human judgment as to whether it is a fit for purpose. And I believe we're up to this, I believe 10 years from now, we will develop the standards process and the governance techniques and the oversight mechanisms, but there's a lot of work ahead particularly in the United States to reach that outcome.

**[00:59:54.8] Thomas Donnelly:** Thank you, mark and Mary Anne Franks the the final word to you.

**[01:00:00.2] Mary Anne Franks:** Thanks so much, and I'll make it quick. The optimistic take would be a lot of the insights from this panel, taking it seriously that this is an opportunity for us to rethink some of our possibly outdated views of the First Amendment and free speech. To think less about the question of, well, what is permitted? And think more about what do we actively want to promote as a society? What do we want to cultivate? What are the values that we care about? What are the kinds of speech and expression that we want to rise to the top? And what are the ways that we can do everything we can to facilitate that and not think so much in terms of the legalistic notions of these issues, but really try to think about ourselves as a collective society, trying to do what is best for each of our citizens and for the achievement of democracy at long last.

**[01:00:45.5] Thomas Donnelly:** Mary Anne Franks, Marc Rotenberg. Mark Coeckelbergh, thank you so much for a wonderful conversation.

**[01:00:55.5] Lana Ulrich:** This program was live streamed on May 15th, 2024. This episode was produced by Tanaya Tauber, Bill Pollock, and me, Lana Ulrich. It was engineered by Greg Scheckler and Bill Pollock. Research was provided by Saman Mostashari, Cooper Smith and Yolo Tyrese. Check out our full lineup of exciting programs and register to join us virtually at [constitutioncenter.org](https://constitutioncenter.org). We'll publish these programs on the podcast, so stay tuned here as well, or watch our videos available in our media library at [constitutioncenter.org/medialibrary](https://constitutioncenter.org/medialibrary). Please rate, review, and subscribe to live at the National Constitution Center on Apple Podcasts, or follow us on Spotify. On behalf of the National Constitution Center, I'm Lana Ulrich.