



The AI app Grok.

Press Association via AP Images

SCIENCE & TECH

How AI deepfakes have skirted revenge porn laws

Limits unclear when explicit images look real, but are digitally generated

Christina Pazzanese | Harvard Staff Writer

January 28, 2026 • 8 min read

Federal and state governments have outlawed “revenge porn,” the nonconsensual online sharing of sexual images of individuals, often by former partners. Last year, South Carolina became the 50th state to enact such a law.

The recent rise of easy-to-use generative AI tools, however, has introduced a new wrinkle: What happens when those images look real but have been created by AI? What’s lawful in the U.S. and who’s responsible is not yet clear.

The U.K. has threatened to ban X and Grok, Elon Musk’s social media platform and AI model, for allowing users to create explicit deepfake images of people, including children.

The European Union, China, and India strictly regulate all AI-generated content. Last week, South Korea passed a sweeping new law that holds GenAI models responsible for their misuse.

The U.S. Senate **passed** a bill this month giving deepfake victims the right to sue; the House has not taken up the measure.

In this edited conversation, **Rebecca Tushnet**, Frank Stanton Professor of the First Amendment at **Harvard Law School** and faculty co-director of the Berkman Klein Center for Internet and Society, explains the legal landscape around the use of AI deepfake technology and the difficulties it presents for U.S. regulators.

What's different about AI deepfakes?

The issue of people fantasizing about other people with their clothes off is, of course, pretty ancient. But historically, the methods for producing and distributing realistic representations of that were quite limited. Sometimes a difference in degree can be a difference in kind, and this might well be one of those times.

The scale is very different, and the realism is very different. The extra realism of AI-generated representations does seem to make a difference to both the injury and intrusion that people feel, and, potentially, the effects on the viewer.

Is creating deepfakes easy across all AI models?

None of the other big models let you do this.

It's possible for many people with some computer science knowledge to take a perfectly ordinary, well-behaved model and then train it with a bunch of porn. It's relatively easy for people who, two years ago, couldn't have done this now to do this.

So even if Grok starts behaving like the other major models, it will still be plenty possible for these images to get generated. It just won't be as easy.

The other thing people are doing is things like "show Billie Eilish making a Nazi salute." It doesn't have the sexual explicitness constraint, and it might not, in some sense, fool anybody.

That makes it really hard, at least in the U.S., to regulate because our regulations are much better at falsity. If you're going to deceive someone into thinking Billie Eilish is making a Nazi salute, that's a problem.

But most people who see these will say, this is a manipulated image. That's really hard to fit into our tradition. Because usually, if you didn't do it to her (or use the image in an ad) we don't regulate that.

It is extremely hard, and especially in the current political context, extremely dangerous to give the government the power to decide which images are manipulated in a bad way.

"In the U.S., the baseline responsibility is to remove intimate depictions when you're informed about them."

— Rebecca Tushnet

What does U.S. law allow and disallow currently?

It's really a big mess. With nudes, there's a little firmer ground to say that there are potential liabilities coming for false light or invasion of privacy. But that's for the person who disseminated the image.

The tool that they use to generate a specific image — how to regulate it is not clear.

There are a lot of states that have laws against unfair and deceptive business conduct, and there's a case to be made that having this capacity is unfair, but it's actually a hard one, because usually when we think about unfairness, we think about unfairness to people as consumers, not unfairness to people in their role as people or children.

In the U.S., the baseline responsibility is to remove intimate depictions when you're informed about them.

The question of liability for being able to generate them is one we just don't have much experience with. It was possible to do a whole bunch of this in previous generations of digital tools. It's just that the tools were largely in the hands of people who didn't want to do that.

In general, creating these depictions of celebrities, but also non-celebrities, is just not a way to make a ton of money. In general (as a society), we haven't said the mere fact that you can use a tool to do this makes the tool unlawful.

But this is new territory and it may be that given where we are, you have a duty to, at least, make it difficult.

With these tools, there are often ways around them. And from a legal perspective, the usual question is: What steps does one have to take, if any, to make it difficult — what they call in AI, "guardrails."

I think there's a wide range of potential legal responses, from "It's only the person who uses the tool who's responsible for what they do," like what we do with guns, all the way to "The toolmaker is an insurer, that is, they are automatically liable if the tool is misused, no matter what they did to try to make the tool hard to use."

And although you see people occasionally suggesting that last option, that's not a usual way that we configure liability. We're still in the process of thinking through exactly what should be required of a toolmaker. Right now, there's certainly no law that even requires guardrails.

What are the primary arguments for and against regulation?

The people who want to make it legally very risky to generate images like this generally focus on human dignity and the disparate effects of this technology, especially on women and minoritized populations, because they're used to degrade in extremely gendered ways. They are used as a tool of supremacy. I think that's a powerful argument.

The people who say that we need to be very careful here point out that, for example, sexualized mockery is a common means of political speech. Before the French Revolution, a whole bunch of pornography circulated about Marie Antoinette. It was both pornographic and political.

And so, they say we want to be extremely cautious not to suppress actual instances of speech or imagination. Something you do on your own time and that you do privately probably shouldn't be a concern of the government. It's only when you distribute it that we have a problem.

The EU and the U.K. regulate the internet and AI robustly. Are they out front on this or is everyone still scrambling to figure it out?

I would say we're all scrambling from very different bases. The American First Amendment is just more robust than most other countries' free speech protection. But we're all scrambling in part because a lot of this stuff probably isn't within the scope of the First Amendment, so our differences don't totally matter.

The thing that makes it distinctively hard for the U.S. is not the First Amendment, but our larger legal culture that has large pockets where it doesn't recognize good-faith attempts.

In Europe in particular, they tend to have very sweeping regulations. But if the regulator says, "You're really doing your best, and we think that you are trying to comply," the regulator will say you are in compliance.

In the U.S., we tend to only do that if the law explicitly says you have to make good-faith efforts to do X. So, we both tend to pass laws saying you have to do X. But in Europe, the regulator thinks that your good-faith efforts count.

In the U.S., private plaintiffs in particular will say, "Oh, the law says you have to do X. Here's a place where you didn't. You are liable." What I've described as making you an insurer. It is a mess.



HARVARD
UNIVERSITY