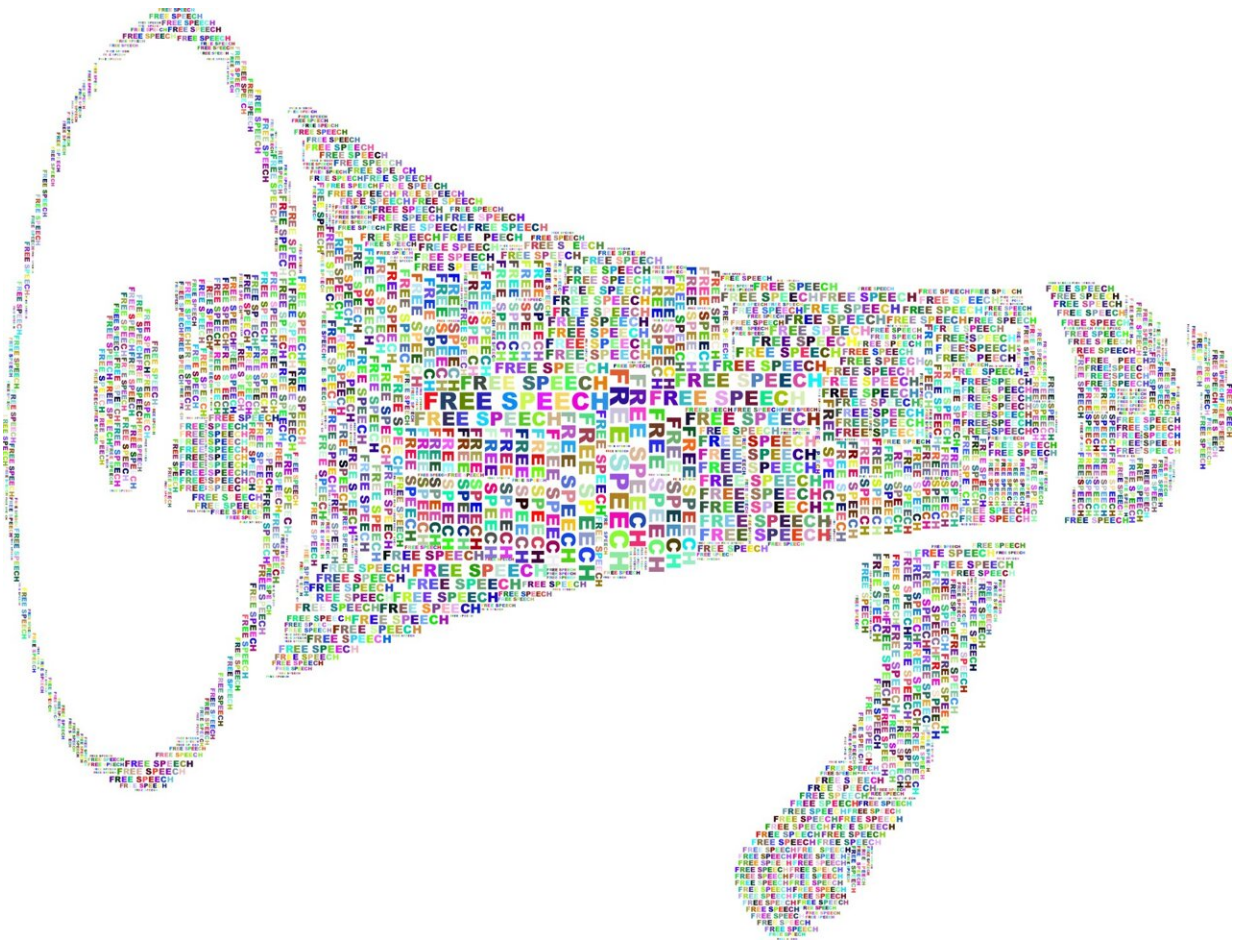


# How AI threatens free speech, and what must be done about it

January 19 2024, by Philip Seargeant

---



Credit: Pixabay/CC0 Public Domain

Headlines about the threats of artificial intelligence (AI) tend to be full

of [killer robots](#), or fears that when they're not on killing sprees, these same robots will be [hoovering up human jobs](#). But a serious danger which gets surprisingly little media attention is the impact these new technologies are likely to have on freedom of expression. And, in particular, how they're able to undermine some of the most foundational legal tenets that protect free speech.

Every time a new communications technology sweeps through society, it disrupts the balance that has previously been struck between social stability and [individual liberty](#).

We're currently living through this. Social media has made new forms of community networking, surveillance and public exposure possible, which have led to increased [political polarization](#), the rise of global populism and an epidemic of [online harassment and bullying](#).

Amid all this, [free speech](#) has become a totemic issue in [the culture wars](#), with its status both boosted and threatened by the societal forces unleashed by [social media platforms](#).

Yet free speech debates tend to be caught up with arguments about "cancel culture" and the "woke" mindset. This risks overlooking the impact technology is having on how freedom of expression laws actually work.

In particular, the way that AI gives governments and tech companies the ability to censor expression with increasing ease, and at great scale and speed. This is a serious issue that I explore in my new book, [The Future of Language](#).

## **The delicate balance of free speech**

Some of the most important protections for [free speech in liberal](#)

[democracies](#) such as the UK and the US rely on technicalities in how the law responds to the real-life actions of everyday citizens.

A key element of the current system relies on the fact that we, as autonomous individuals, have the unique ability to transform our ideas into words and communicate these to others. This may seem a rather unremarkable point. But the way the law currently works is based on this simple assumption about human social behavior, and it's something that AI threatens to undermine.

Free speech protections in many liberal societies rule against the use of "[prior restraint](#)"—that is, blocking an utterance before it's been expressed.

The government, for instance, should not be able to prevent a newspaper from publishing a particular story, although it can prosecute it for doing so after publication if it thinks the story is breaking any laws. The use of prior restraint is already widespread in countries such as [China](#), which have very different attitudes to the regulation of expression.

This is significant because, despite what tech libertarians such as [Elon Musk may assert](#), no society in the world allows for absolute freedom of speech. There's always a balance to be struck between protecting people from the real harm that language can cause (for example by defaming them), and safeguarding people's right to express conflicting opinions and criticize those in power. Finding the right balance between these is one of the most challenging decisions a society is faced with.

## **AI and prior restraint**

Given that so much of our communication today is mediated by technology, it is now extremely easy for AI assistance to be used to enact prior restraint, and to do so at great speed and massive scale. This would

create circumstances in which that basic human ability to turn ideas into speech could be compromised, as and when a government (or social media exec) wishes it to be.

The UK's recent [Online Safety Act](#), for instance, as well plans in the US and Europe to use "[upload filtering](#)" (algorithmic tools for blocking certain content from being uploaded) as a way of screening for offensive or illegal posts, all encourage social media platforms to use AI to censor at source.

The rationale given for this is a practical one. With such a huge quantity of content being uploaded every minute of every day, it becomes extremely challenging for teams of humans to monitor everything. AI is a fast and far less expensive alternative.

But it's also automated, unable to bring real-life experience to bear, and its decisions are rarely subject to public scrutiny. The consequences of this are that AI-driven filters can often lean towards censoring content which is neither illegal or offensive.

Free speech as we understand it today relies on specific legal processes of protection that have developed over centuries. It's not an abstract idea, but one grounded in very particular social and legal practices.

Legislation that encourages content regulation by automation effectively dismisses these processes as technicalities. In doing so, it risks jeopardizing the entire institution of free speech.

Free [speech](#) will always be an idea sustained by ongoing debate. There's never a settled formula for defining what should be outlawed and what not. This is why determining what counts as acceptable and unacceptable needs to take place in open society and be subject to appeal.

While there are indications that some governments are beginning to [acknowledge this](#) in planning for the future of AI, it needs to be center stage in all such plans.

Whatever role AI may play in helping to monitor online content, it mustn't constrain our ability to argue among ourselves about what sort of society we're trying to create.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation

Citation: How AI threatens free speech, and what must be done about it (2024, January 19)  
retrieved 2 May 2024 from <https://techxplore.com/news/2024-01-ai-threatens-free-speech.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.