



THE
FUTURE
OF
FREE
SPEECH

THAT VIOLATES MY POLICIES 2026

AI LAWS, CHATBOTS, AND
THE FUTURE OF EXPRESSION

DIRECTED BY:

Jordi Calvet-Bademunt,
Jacob Mchangama,
and Isabelle Anzabi

www.futurefreespeech.org

ACKNOWLEDGMENTS

The Future of Free Speech is an independent, nonpartisan think tank based at Vanderbilt University. Our mission is to reaffirm freedom of expression as the foundation of free and thriving societies through actionable research, practical tools, and principled advocacy. We envision a world in which the right to freedom of expression is safeguarded by law and strengthened by a culture that embraces diverse viewpoints.

This project was led by Jordi Calvet-Bademunt (Senior Research Fellow), Jacob Mchangama (Executive Director), and Isabelle Anzabi (Research Associate) at The Future of Free Speech. Jordi Calvet-Bademunt authored the chapter on the European Union, and Isabelle Anzabi and Ashkhen Kazaryan (Senior Legal Fellow) authored the chapter on the United States.

We are grateful to Ajdin Muratovic (Director of Communications) for overseeing the report's design, and to Wendy H. Burch (Chief Operating Officer) for coordinating all administrative aspects of the project.

We extend our thanks to the leading experts who contributed chapters on their respective jurisdictions: Tomiwa Ilori (Nigeria); Matías González Mama, Lina Palacios, and Francesca Chocano Villanueva (Peru); Emmanuel Vargas Penagos, Katalin Kelemen, and Sandi Rizvic (Sweden); Yaman Akdeniz (Türkiye); Eliza Bechtold (United Kingdom); and Trịnh Hữu Long (Vietnam).

We are further indebted to Justin Hayes for designing the report.

Finally, we are especially grateful to the Rising Tide Foundation and the Swedish Postcode Lottery Foundation for their generous support of this work.



PREFACE

In the 2026 edition of our report *“That Violates My Policies”: AI Laws, Chatbots, and the Future of Expression*, we explore the ways in which public and private governance of generative artificial intelligence (AI) shapes the space for free expression and access to information.

Since the launch of ChatGPT by OpenAI in November 2022, generative AI has moved from a novelty to a tool embedded in everyday communication and information access.¹ As its reach has grown, so has the urgency for legislators around the world to determine how it should be governed, and what that governance means for freedom of expression and access to information.

Recognizing generative AI’s growing impact on freedom of expression and access to information, The Future of Free Speech undertook a first-of-its-kind analysis of how national legislation and corporate practices shape expression in the era of generative AI. The 2025 report analyzed AI legislation in Brazil, China, the European Union, India, the Republic of Korea, and the United States. It focused on laws and public policies addressing AI-generated content, particularly around elections and political speech, hate speech, defamation, explicit content, copyright, and measures that actively promote freedom of expression.²

This year, we built on those findings and expanded them considerably. This report, the first of two publications The Future of Free Speech will release this year, examines AI legislation, bills, policies, and case law in the European Union, Nigeria, Peru, Sweden, Türkiye, the United Kingdom, the United States, and Vietnam. It considers how these frameworks shape

¹ Generative AI refers to systems that create content — including text, images, video, audio, and software code — in response to user prompts. Cole Stryker and Mark Scapicchio, “What Is Generative AI?,” IBM Think, March 22, 2024, <https://www.ibm.com/think/topics/generative-ai>.

² Jordi Calvet-Bademunt, Jacob Mchangama, and Isabelle Anzabi, *“That Violates My Policies”: AI Laws, Chatbots, and The Future of Expression* (Nashville: The Future of Free Speech, October 2025), <https://futurefreespeech.org/wp-content/uploads/2025/10/AI-Report-2025-Full-Report.pdf>.

THAT VIOLATES MY POLICIES 2026

freedom of expression and access to information across election and political content, hate speech, defamation, explicit content, copyright, and measures that empower freedom of expression, such as policies supporting linguistic diversity, media literacy, and open-source development. We have added dedicated sections on protection of minors and AI-powered surveillance to reflect these increasingly salient issues.

This report seeks to provide a rigorous and timely analysis of how generative AI is reshaping the space for free expression in both the public and private spheres. Building on these insights, The Future of Free Speech is developing guidelines to help policymakers ensure that generative AI protects and enhances freedom of expression and access to information, two cornerstones of democratic societies. A second report, focused on model analysis and related good practices for companies, will follow later this year.

In an era of rapid technological change, safeguarding free expression protects both individual rights and the conditions for open, informed, and thriving democracies.

A decorative graphic in the top-left corner consisting of a network of blue dots connected by thin lines, resembling a molecular or digital structure.

EXECUTIVE SUMMARY

Do laws governing AI safeguard or constrain freedom of expression? Our report, *“That Violates My Policies”: AI Laws, Chatbots, and the Future of Expression (2026 Edition)*, addresses this question by assessing legislation, public policies, and court decisions in eight jurisdictions: the European Union, Nigeria, Peru, Sweden, Türkiye, the United Kingdom, the United States, and Vietnam. Here we provide the report’s executive summary.

Key Findings

Several trends stand out across the eight jurisdictions examined in the 2026 edition of this report. Most restrictions on AI-generated expression do not result from new AI laws. Rather, they arise from existing rules on cybercrime, defamation and insult, hate speech, terrorism, and online safety that apply, or will apply, to AI-generated content. Where AI-specific rules exist, they often rely on broad concepts such as “systemic risk,” “national security,” “public order,” or “neutrality,” leaving wide discretion to regulators, prosecutors, and companies. In some countries, short takedown deadlines and heavy sanctions create strong incentives to over-remove content, including lawful content. Enforcement is already underway, including court-ordered blocking of chatbot posts in Türkiye and prison sentences in Vietnam for producing AI-generated content. Several jurisdictions are also expanding AI-enabled surveillance, including facial recognition at protests, which risks chilling assembly and dissent. Minority-language AI and media literacy initiatives offer more rights-protective approaches, but many commitments remain on paper. Policymakers should recommit to protecting the freedoms of expression, access to information, and thought; any restrictions must be clearly defined by law, pursue a legitimate aim, and be proportionate.

European Union

The EU adopted the world's first cross-sector AI framework, the AI Act, and several of its provisions started to bite in 2026. On paper, the EU has strong protections for free expression, and the Charter of Fundamental Rights and the European Convention on Human Rights guarantee the right to receive and share information. Some AI Act provisions protect expression directly, including limits on biometric surveillance by police. In practice, however, Europe has been too willing to restrict contentious speech, particularly through vague definitions of hate speech, insult, and even disinformation, and courts have often allowed governments and institutions wide latitude to restrict expression and access to information. Vaguely worded systemic-risk obligations in the AI Act and the Digital Services Act, which now also covers ChatGPT, invite misguided enforcement and self-censorship. Applied with genuine commitment to freedom and pluralism, the Charter and the Convention offer the guidance needed to protect expression. Restrictions must be precise, pursue a legitimate aim, and be necessary and proportionate, adopted only after considering less restrictive alternatives.

Nigeria

In Nigeria, free expression is guaranteed by the constitution and by regional human rights agreements but is often narrowed in practice. Cybercrime, criminal, and broadcasting laws use vague language that leaves room for abuse, and a regional court has found some provisions to violate free expression rights. A bill before the legislature would create a dedicated AI regulator, and a national AI strategy sets broad goals, but neither addresses free expression in AI. As a result, harms such as nonconsensual intimate imagery can scale largely unchecked, while the same vague laws raise the risk of overbroad enforcement against journalists and critics, encouraging self-censorship on politically sensitive topics. Nigeria already has the constitutional and regional foundation that could guide rights-respecting AI regulation. Building on it requires an explicit rights-based framework that incorporates transparency and accountability.

Peru

Peru has moved quickly to build a legal framework for artificial intelligence and became the first country in Latin America to adopt a comprehensive AI law, alongside a separate law making the use of AI an aggravating factor in criminal offenses. These advances coexist with several tensions. Rules ban harmful uses of AI, including mass surveillance “without a legal basis,” but lack an autonomous sanctions regime, leaving protections largely symbolic. Real-time facial recognition in public is formally banned, yet camera networks have expanded quickly and have already been used against protesters. AI-specific criminal provisions include no exception for satire or parody and could amplify chilling effects against journalists and critics. Peru’s 2026 election proceeded without a law addressing electoral deepfakes, despite hundreds of disinformation alerts. Indigenous-language AI initiatives support information access. If Peru can pair its regulatory ambition with institutional capacity for enforcement, its framework could serve as a model in the region.

Sweden

Sweden has a long-standing constitutional commitment to press freedom and free expression, reflected in its four fundamental laws. Beyond the EU AI Act, Sweden has not adopted a framework specifically regulating generative AI, and Swedish courts have not yet addressed it directly as a free expression issue. The national AI strategy prioritizes development, digital sovereignty, literacy, and resilience over new restrictions. Existing rules on political speech, hate speech, defamation, and nonconsensual intimate imagery, among others, may apply to AI-generated material, with liability generally falling on whoever publishes or distributes it rather than on the AI system. AI is increasingly relevant to state surveillance. A 2026 reform permits police to use real-time facial recognition, and the government treats AI as a strategic resource for national security.

Türkiye

Türkiye is preparing for the age of artificial intelligence while dismantling the digital rights infrastructure that would protect freedom of expression. It has no AI-specific legislation, though a commission recommended adopting a framework in early 2026, and it still has not signed the related Council of Europe convention. The deeper constraint on expression is

EXECUTIVE SUMMARY

an existing censorship system now being extended to AI. Authorities have blocked roughly 1.6 million websites and domains, and in 2025 courts began invoking those same powers against AI chatbots—among the earliest known instances of state censorship of generative AI. Courts have treated chatbot output as functionally equivalent to journalism. Insult offenses, especially speech directed at the president, already produce tens of thousands of investigations a year. A new law taking effect in late 2026 will require takedowns within one hour and automated filtering of removed content. Türkiye’s eventual AI legislation will likely extend the logic of its current censorial digital governance.

United Kingdom

The Labour government has promised AI-specific legislation, but for now the UK regulates AI only by extending existing rules, chiefly the Online Safety Act. Free expression is protected through the Human Rights Act and the UK’s membership in the European human rights system. The Online Safety Act’s severe penalties push platforms toward removing content on mere suspicion of illegality, and the UK’s broad hate speech and terrorism laws could just as easily reach AI-generated political speech. A 2026 law lets ministers extend the Act’s duties to AI-generated content. No comprehensive law governs police use of AI-powered surveillance tools. New offenses now target nonconsensual explicit deepfakes. The devolved governments run their own media literacy programs aimed at improving accessibility. The government is investing billions in AI without increasing funding for the independent body that protects human rights—a gap that leaves free expression poorly positioned as the UK’s AI plans advance.

United States

Generative AI is testing the United States’ established free speech protections, and courts—rather than Congress—have emerged as the key check. First Amendment rulings have struck down state election-deepfake laws in California and Hawaii for sweeping in too much protected speech, while allowing product-liability claims against an AI companion chatbot company to proceed. There is still no comprehensive federal AI law. Instead, a federal preemption campaign, pursued through executive orders and litigation, has narrowed state AI governance without establishing a statutory framework in its place. A new federal law targets AI-generated nonconsensual explicit content, though civil liberties groups

EXECUTIVE SUMMARY

warn it could pressure platforms into removing lawful content. Federal efforts to push “ideological neutrality” in AI outputs raise concerns about government pressure on speech. AI-powered government surveillance, including facial recognition at protests, is chilling First Amendment rights. Any regulatory response should stay narrowly focused to preserve the country’s tradition of speech protection.

Vietnam

Vietnam’s constitution and ratified treaties guarantee free expression but carry little weight in what is widely regarded as a rule-by-law system without an independent judiciary to enforce them. Vietnam passed its first AI law in 2025, alongside an updated cybersecurity law, both using vague language to ban AI-generated content deemed a threat to national security or public order. These broad terms give law enforcement wide discretion to apply them to criticism of the regime or government officials. Two exiled dissidents were sentenced in absentia to seventeen years in prison, partly for AI-generated videos of leaders, and fines have followed AI-generated fake news. A national identity app used by most of the population is to be integrated with AI, drawing on citizens’ communications data, with no independent oversight. Provisions on minority language support are largely symbolic, and AI regulation in Vietnam functions mainly to extend the state’s existing control over speech.



**THE
FUTURE
OF
FREE
SPEECH**

www.futurefreespeech.org