



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>.

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.



ARTIFICIAL INTELLIGENCE & FREEDOM OF EXPRESSION IN THE UNITED STATES

By

Isabelle Anzabi & Ashkhen Kazaryan



ISABELLE ANZABI

Isabelle Anzabi is a research associate at The Future of Free Speech at Vanderbilt University, where she analyzes the intersections between AI policy and freedom of expression. Isabelle is also President of the Internet Law & Policy Foundry. Previously, she was an AI & Human Rights Fellow with the European Center for Not-for-Profit Law and a Knowledge Fellow at the DiploFoundation. Isabelle received her B.A. in Political Science with a minor in Psychology from Stanford University. She studied digital governance at Oxford University and interned at institutions including the World Bank and CISA. At Stanford, she was affiliated with the Stanford Center for Racial Justice, Legal Design Lab, Cyber Policy Center, Constitutional Law Center, Stanford Technology Law Review, and Public Service Leadership Program.



ASHKHEN KAZARYAN

Ashkhen Kazaryan is a nationally recognized expert in First Amendment law and technology policy, specializing in digital free speech, artificial intelligence, and the intersection of constitutional rights with emerging technologies. As Senior Legal Fellow at The Future of Free Speech at Vanderbilt University, she leads legal, policy, and advocacy initiatives addressing some of the most consequential questions affecting expression in the United States. Her work includes legal and policy analysis, regulatory advocacy, coalition building, policymaker engagement and public commentary. Previously, Ashkhen was the lead for North and Latin America on the content regulation team at Meta, where she also served as the company's policy lead on Section 230.

Authors' Note: We thank Jeff Kosseff for his valuable comments and suggestions.

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

ABSTRACT

The continued rise in popularity and impact of generative artificial intelligence (AI) in the United States is testing established free speech protections. Courts have emerged as the key check on AI speech regulation, with First Amendment rulings striking down state election deepfake laws in California and Hawaii and allowing product liability claims against AI chatbot companies to proceed on design-defect theories. Despite the absence of comprehensive federal AI legislation, the United States retains stronger constitutional protections for AI-generated speech than most other jurisdictions. Yet those protections are increasingly tested from within. A federal preemption campaign, pursued by some members of Congress, as well as advanced through executive orders, litigation task forces, and funding conditions, has narrowed or supplanted state AI governance without replacing it with a statutory framework. New technological developments have expanded AI-powered government surveillance, which in turn has chilled First Amendment rights. Examining the application of these laws and principles, this chapter demonstrates that any regulatory response must narrowly target concrete, demonstrable harms, preserving the constitutional principles that make the United States an outlier in protecting expressive freedom.

1. INTRODUCTION

The United States has one of the most robust systems of free speech protection in the world, anchored in the First Amendment command that “Congress shall make no law ... abridging the freedom of speech, or of the press.”¹ The Fourteenth Amendment incorporates this right to the states, and the First Amendment also extends these protections to all three branches of government—executive, judicial, and legislative—at both the federal and state levels.² The Supreme Court has interpreted this protection broadly, extending it to new communication technologies and safeguarding both the right to speak and the right to receive information and ideas.³ Generative AI presents the next frontier that courts need to explore.

Generative AI’s capability to produce text, images, audio, and video at scale offers unprecedented opportunities to expand access to information and lower barriers to participation in public debate. It can serve as a creative and educational tool, a means of preserving cultural and linguistic diversity, and a way to make information more accessible to people. Yet the same capabilities raise novel questions about liability, truthfulness, and the potential for misuse.

Much of the regulatory activity has occurred at the state level. Key areas of state legislative activity have included algorithmic accountability, AI use in hiring and employment, protection of minors, and expanded protections against nonconsensual intimate imagery (NCII). Courts are the forums where legal theories are tested against new technology. They have struck down several state AI laws on First Amendment grounds, allowed product liability claims against AI chatbot companies to proceed, and clarified the boundaries of copyright protection in AI training.

¹ U.S. Const. amend. I.

² Congressional Research Service, “State Action Doctrine and Free Speech,” *Constitution Annotated*, https://constitution.congress.gov/browse/essay/amdt1-7-2-4/ALDE_00013541/.

³ *Reno v. American Civil Liberties Union*, 521 U.S. 844, 868–70 (1997); *Stanley v. Georgia*, 394 U.S. 557, 564 (1969).

Against that judicial backdrop, executive action in late 2025 and 2026 introduced a push towards preempting state AI governance, conditioning federal funding on states' AI regulatory choices, and deploying a new DOJ litigation task force to challenge state laws directly.⁴ Meanwhile at the state level, Colorado's landmark AI Act was repealed and replaced with a narrower statute, while California's and Hawaii's deepfake election laws were struck down on First Amendment grounds in federal court. Additionally, a \$1.5 billion copyright settlement established new limits on how AI companies may acquire training data. Finally, AI-powered federal surveillance has chilled expression and assembly.

This chapter examines these emerging challenges in detail, examining federal and state regulatory trends, the unsettled question of AI's status under the First Amendment, and the constitutional limits on regulating harmful or deceptive content. While acknowledging legitimate concerns about AI's potential misuse, it argues that the United States should address these risks in ways that preserve the country's long-standing commitment to the First Amendment.

⁴ Exec. Order No. 14365, 90 Fed. Reg. 58499 (Dec. 16, 2025), <https://www.federalregister.gov/documents/2025/12/16/2025-23092/ensuring-a-national-policy-framework-for-artificial-intelligence>; The White House, "National Legislative Policy Framework for Artificial Intelligence," March 20, 2026, <https://www.whitehouse.gov/releases/2026/03/president-donald-j-trump-unveils-national-ai-legislative-framework/>.

2. GENERAL STANDARDS OF FREEDOM OF EXPRESSION

The First Amendment explicitly prohibits Congress from enacting any law that abridges the freedom of speech and, more broadly, limits government restrictions on the dissemination and receipt of information and ideas. It also protects access to information and ideas.⁵ Importantly, it also protects the right to receive information and ideas “regardless of their social worth,” underscoring the First Amendment’s protection against government-imposed viewpoint discrimination.⁶

The Supreme Court has established a framework wherein different categories of speech receive varying levels of protection under the First Amendment.⁷ Political, ideological, and artistic speech lie at the core of First Amendment protection, while commercial speech receives less protection.⁸ The court has also identified categories of speech that fall outside First Amendment protection, including incitement to imminent lawless action, true threats, fraud, defamation, obscenity, and child pornography (also referred to as child sexual abuse material or CSAM). Determining the constitutionality of regulations on protected speech hinges on whether the regulation is content-based or content-neutral.⁹ Content-based restrictions are not automatically unconstitutional, but they are subject to strict scrutiny, which is an exceptionally high standard that requires the government to demonstrate that the law is the least restrictive means of advancing a compelling governmental interest.¹⁰ In contrast, content-neutral restrictions, such as time, place, and manner regulations, are evaluated under intermediate scrutiny, which requires the law to be narrowly tailored to

5 *Lamont v. Postmaster General*, 381 U.S. 301 (1965).

6 *Stanley v. Georgia*, 394 U.S. 557 (1969).

7 Congressional Research Service, *The First Amendment: Categories of Speech*, CRS In Focus IF11072 (March 28, 2024), <https://www.congress.gov/crs-product/IF11072>.

8 *Central Hudson Gas & Elec. v. Public Svc. Comm’n*, 447 U.S. 557 (1980).

9 David L. Hudson Jr., “Content Based,” *First Amendment Encyclopedia*, August 10, 2023, <https://firstamendment.mtsu.edu/article/content-based>; David L. Hudson Jr., “Content Neutral,” The Free Speech Center, January 1, 2009, <https://firstamendment.mtsu.edu/article/content-neutral>; “Overview of Content-Based and Content-Neutral Regulation of Speech,” *Constitution Annotated*, Library of Congress, accessed August 1, 2025, https://constitution.congress.gov/browse/essay/amdt1-7-3-1/ALDE_00013695/.

10 *United States v. Playboy Entertainment Group, Inc.*, 529 U.S. 803, 813 (2000).

serve a substantial governmental interest.¹¹ Commercial speech also receives a form of intermediate scrutiny, although the applicable analysis differs depending on the nature of the regulation.¹² The court has held that viewpoint discrimination is the most “egregious” form of content discrimination.¹³

2.1 Editorial Discretion and Private Actors

The Supreme Court has also consistently protected editorial rights across a broad spectrum of diverse media. The landmark *Miami Herald Publishing Co. v. Tornillo* was pivotal in defining the extent of editorial discretion under the First Amendment. The Supreme Court upheld a newspaper’s autonomy over its content choices and treatment of public issues and officials, holding that such editorial choices, “whether fair or unfair, constitute the exercise of editorial control and judgment.”¹⁴ Later, the Supreme Court made it clear that editorial protections established in *Tornillo* are not limited to newspapers. In *Pacific Gas & Electric Co. v. Public Utilities Commission of California*, the court struck down a regulatory rule that compelled a utility company to distribute political editorials critical of the company within its billing envelopes. Time and again, the Supreme Court has reaffirmed that editorial discretion extends beyond newspapers and other forms of “traditional media.” In *Hurley v. Irish-American Gay, Lesbian and Bisexual Group of Boston*, a state court had previously determined that, under public accommodation law, a private parade organized by an association must include a gay rights advocacy group. The parade’s organizers contested this, stating that including the gay rights group would contradict the message they intended to convey. Recently, in *Moody v. NetChoice*, the Supreme Court explained that social media platforms exercise protected editorial discretion under the First Amendment when deciding whether to publish, remove, promote, or demote user-generated content through algorithmic systems.¹⁵ Taken together, these decisions show that the Supreme Court jurisprudence has expanded the First Amendment protections to new technologies.¹⁶

¹¹ *Ward v. Rock Against Racism*, 491 U.S. 781, 791, 798–800 (1989).

¹² *Central Hudson Gas & Electric Corp. v. Public Service Commission of New York*, 447 U.S. 557, 563–66 (1980); *Zauderer v. Office of Disciplinary Counsel of the Supreme Court of Ohio*, 471 U.S. 626, 650–51 (1985).

¹³ “Overview of Viewpoint-Based Regulation of Speech,” *Constitutional Annotated*, Legal Information Institute, accessed August 1, 2025, <https://www.law.cornell.edu/constitution-conan/amendment-1/overview-of-viewpoint-based-regulation-of-speech>.

¹⁴ *Ibid.*

¹⁵ *Moody v. NetChoice, LLC*, 144 S.Ct. 2383 (2024).

¹⁶ *Brown et al. v. Entertainment Merchants Ass’n et al.*, 564 U.S. 786 (2011).

2.2 Applying First Amendment Principles to Generative AI

The application of freedom of expression principles to AI-generated content is the subject of ongoing legal battles. US legislators and courts face many distinct legal questions, the first one being whether AI-generated expression constitutes “speech” under the First Amendment. Generative AI is a tool for creating expressive content, and similar to the printing press or cameras, it “make[s] it easier to speak.”¹⁷ Scholars have argued that the First Amendment protects the rights of creators and users, and restricting AI-generated content could infringe on their right to express themselves.¹⁸

The academic and policy focus has been on whether AI systems can themselves be speakers with protected output, or whether only the humans who create, deploy, or interact with AI systems have legal rights.¹⁹ Some argue against full First Amendment protections for AI output, viewing it as lacking the human intentionality that traditionally underlies free speech rights.²⁰

Another issue that legal scholars raise is that, in the context of AI-generated speech, the focus should be on the listener’s right to receive information, regardless of whether the source is human or artificial.²¹ The Supreme Court recognizes the right to receive information as a corollary to the right to speak, supporting the view that users have a right to obtain information from AI models.²² Some scholars have suggested that even AI output generated without human intervention should be protected.²³

Lastly, developers and AI platforms may assert that the design and code of their systems

17 Eugene Volokh, “First Amendment Limits on AI Liability,” *Lawfare*, September 27, 2024, <https://www.lawfaremedia.org/article/first-amendment-limits-on-ai-liability>.

18 *Ibid.*

19 Eugene Volokh, Mark A. Lemley, and Peter Henderson, “Freedom of Speech and AI Output,” *Journal of Free Speech Law* 3 (August 3, 2023): 651, <https://www.journaloffreespeechlaw.org/volokhlemleyhenderson.pdf>; “AI and the First Amendment: A Q&A with Jack Balkin,” Yale Law School, January 29, 2024, <https://law.yale.edu/yls-today/news/ai-and-first-amendment-qa-jack-balkin>.

20 Peter Salib, “AI Outputs Are Not Protected Speech,” *Washington University Law Review* 102, University of Houston Law Center Research Paper no. 2024-A-5, January 1, 2024, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4687558; Karl M. Manheim and Jeffery Atik, “AI Outputs and the Limited Reach of the First Amendment,” *Washburn Law Journal* 63 (2024): 159, <https://ssrn.com/abstract=4626235>; *Garcia v. Character Technologies*, 785 F. Supp. 3d 1157, 1179-80.

21 Eugene Volokh, “First Amendment Limits on AI Liability,” *Lawfare*, September 27, 2024, <https://www.lawfaremedia.org/article/first-amendment-limits-on-ai-liability>.

22 *Virginia State Board of Pharmacy v. Virginia Citizens Consumer Council, Inc.*, 425 U.S. 748, 756-57 (1976); Eugene Volokh, “First Amendment Limits on AI Liability,” *Lawfare*, September 27, 2024, <https://www.lawfaremedia.org/article/first-amendment-limits-on-ai-liability>.

23 Eugene Volokh, Mark A. Lemley, and Peter Henderson, “Freedom of Speech and AI Output,” *Journal of Free Speech Law* 3 (August 3, 2023): 651, <https://www.journaloffreespeechlaw.org/volokhlemleyhenderson.pdf>.

ARTIFICIAL INTELLIGENCE AND FREEDOM OF EXPRESSION IN THE UNITED STATES

constitute protected speech under the First Amendment. Whether such claims will succeed remains unsettled by law.

3. AI LEGISLATION LANDSCAPE

3.1 The Council of Europe AI Framework Convention

The Council of Europe's Framework Convention on Artificial Intelligence and Human Rights, Democracy, and the Rule of Law, is the first international treaty devoted to AI-related risks. The principles invoked include human dignity, transparency, accountability, equality, privacy, and reliability.²⁴ Importantly, while the United States signed the Convention in 2024, it has not ratified it as of August 2026, which means that the treaty has not entered into force for the United States.²⁵

3.2 Federal AI Policy: Deregulation and Preemption

Over the past couple of years, US federal policy has undergone a marked transformation with a deliberate pivot toward deregulation and innovation. This shift became clear on the first day of President Donald Trump's second administration through the issuance of the executive order titled Initial Rescissions of Harmful Executive Orders and Actions.²⁶ This revoked President Joe Biden's 2023 executive order Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, which had previously guided US federal policy on AI with a focus on ethics and safety.²⁷

A plethora of executive actions followed. The White House Office of Management and Budget revised guidance on federal use and procurement of AI.²⁸ The White House also

24 Council of Europe: Committee of Ministers, "Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law," CETS No. 225, adopted May 17, 2024, <https://rm.coe.int/1680afae3c>.

25 Council of Europe, Treaty Office, Chart of Signatures and Ratifications of Treaty 225, accessed August 31, 2026, <https://www.coe.int/en/web/conventions/full-list?module=signatures-by-treaty&treaty=225>.

26 Exec. Order No. 14148, 90 Fed. Reg. 8237 (January 28, 2025), <https://www.federalregister.gov/documents/2025/01/28/2025-01901/initial-rescissions-of-harmful-executive-orders-and-actions>.

27 Exec. Order No. 14148; Exec. Order No. 14110, 88 Fed. Reg. 75191 (Nov. 1, 2023), <https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence>.

28 Office of Management and Budget, "Accelerating Federal Use of AI through Innovation, Governance, and Public Trust," memorandum M-25-21, April 3, 2025; Office of Management and Budget, "Driving Efficient Acquisition of Artificial Intelligence in Government," memorandum M-25-22, April 3, 2025.

released its America’s AI Action Plan as the centerpiece of the executive branch’s AI policy. The plan endorsed open-weight models, directed federal agencies to procure “unbiased” AI systems to assure “ideological neutrality,” and sought to empower American workers.²⁹

In December 2025, Executive Order 14365 created a Department of Justice task force to challenge state AI laws that the administration deems inconsistent with federal law.³⁰ In March 2026, the White House released its National Policy Framework for AI, which, similar to preceding actions, attempts to reinforce federal supremacy by proposing preemption of state AI regulation, with specific carve-outs for child safety, fraud prevention, and consumer protection.³¹

Despite the administration’s earlier pledge to deregulate AI in order to prioritize innovation, a June 2026 executive order established a voluntary process through which developers may give the federal government up to thirty days of early access to frontier models with advanced cybersecurity capabilities.³² The administration denied that the process amounted to a mandatory preclearance requirement for such models.

In July 2026, the Federal Trade Commission issued a proposed policy statement suggesting that AI companies might violate section 5 of the FTC Act if they steer AI outputs toward ideological or biased objectives.³³ That approach closely aligns with the administration’s AI Action Plan’s focus on neutrality. The statement also echoes a recurring theme—asserting supremacy over state regulation—by claiming that a state law requiring conduct prohibited by section 5 could be impliedly preempted.³⁴ Enforcement based on government judgments about the accuracy or ideological character of disputed answers could raise First Amendment concerns, while coercive regulatory pressure on intermediaries could also implicate *NRA v. Vullo*, a landmark Supreme Court ruling on government jawboning.³⁵ In it

29 The White House, *Winning the Race: America’s AI Action Plan* (July 2025); Exec. Order 14319, 90 Fed. Reg. 35389 (July 28, 2025), <https://www.federalregister.gov/documents/2025/07/28/2025-14217/preventing-woke-ai-in-the-federal-government>.

30 Exec. Order No. 14365, 90 Fed. Reg. 58499 (Dec. 16, 2025), <https://www.federalregister.gov/documents/2025/12/16/2025-23092/ensuring-a-national-policy-framework-for-artificial-intelligence>.

31 The White House, “National Legislative Policy Framework for Artificial Intelligence,” March 20, 2026, <https://www.whitehouse.gov/releases/2026/03/president-donald-j-trump-unveils-national-ai-legislative-framework/>.

32 Exec. Order No. 14409, 91 Fed. Reg. 34565 (June 5, 2026).

33 Federal Trade Commission, “Policy Statement Concerning the Suppression of Accuracy in Artificial Intelligence Systems,” proposed policy statement, 91 Federal Register (July 7, 2026).

34 Ibid.

35 *National Rifle Association of America v. Vullo*, 602 U.S. 175 (2024).

the Court ruled that the government may not use its power to coerce private parties into suppressing speech it disfavors.³⁶

3.3 State AI Legislation: A Growing 50-State Patchwork

In the absence of a unified federal AI framework, states have become the primary actors shaping the legal and normative contours of AI governance. By March 2026, lawmakers in forty-five states had introduced 1,561 AI-related bills.³⁷ By July 1, 2026, the NYU Center on Technology Policy counted 109 enacted laws regulating AI in some form across twenty-nine states. Some researchers attribute this decline from thirty-nine states a year earlier to the chilling effect of federal preemption pressure.³⁸

3.3.1 High-Risk and Specific Uses of AI

States have taken different approaches to AI regulation. For example, Colorado initially adopted a broad framework imposing duties on developers and deployers of “high-risk” AI, but replaced it in 2026 with the narrower Automated Decision-Making Technology Act, centered on notice and transparency.³⁹ In Texas, the Responsible Artificial Intelligence Governance Act, meanwhile, prohibits specific intentional uses of AI in categories such as unlawful discrimination, behavioral manipulation, and interference with constitutional rights, and reserves enforcement to the attorney general.⁴⁰ Connecticut’s law, effective October 2026, reaches further, covering automated employment decisions, AI companions,

36 Ibid.

37 MultiState, “State AI Legislation Tracker 2026: All 50 States,” <https://www.multistate.ai/artificial-intelligence-ai-legislation> (data as of March 2026).

38 Scott Babwah Brennen, “Where State AI Legislation Stands Half Way Into 2026,” Tech Policy Press, July 6, 2026, <https://techpolicy.press/where-state-ai-legislation-stands-half-way-into-2026>.

39 S.B. 24-205, 74th Gen. Assemb. (Colo. 2024); S.B. 26-189, “Concerning the Use of Automated Decision-Making Technology in Sequential Decisions,” 75th Gen. Assemb. (Colo. 2026) (signed May 14, 2026; eff. Jan. 1, 2027), <https://leg.colorado.gov/bills/sb26-189>; xAI LLC v. Weiser, No. 1:26-cv-01515 (D. Colo.) (enforcement stay granted Apr. 27, 2026). S.B. 26-189, “Automated Decision-Making Technology,” 75th Gen. Assemb., 2d Reg. Sess. (Colo. 2026) (signed May 14, 2026; eff. Jan. 1, 2027), <https://leg.colorado.gov/bills/sb26-189>. S.B. 26-189, 75th Gen. Assemb. (Colo. 2026) (signed May 14, 2026; eff. Jan. 1, 2027), <https://leg.colorado.gov/bills/sb26-189>; U.S. Department of Justice, “Justice Department Intervenes in xAI Lawsuit Challenging Colorado’s ‘Algorithmic Discrimination’ Law,” Office of Public Affairs, April 24, 2026, <https://www.justice.gov/opa/pr/justice-department-intervenes-xai-lawsuit-challenging-colorados-algorithmic-discrimination>; xAI LLC v. Weiser, No. 1:26-cv-01515-DDD (D. Colo.); court stay order, Apr. 27, 2026.

40 H.B. 149, “Texas Responsible Artificial Intelligence Governance Act,” 89th Leg., Reg. Sess. (Tex. 2025) (signed June 22, 2025; effective Jan. 1, 2026), <https://capitol.texas.gov/tlodocs/89R/billtext/pdf/HB00149F.pdf>; Jason M. Loring and Graham H. Ryan, “Texas Enacts AI Law Targeting Harmful Use, Fostering Innovation,” National Law Review, June 24, 2025, <https://natlawreview.com/article/texas-enacts-responsible-ai-governance-act>.

synthetic content, whistleblower protection, and state AI infrastructure.⁴¹

3.3.2 AI Content Disclosure: Watermarks and Training Data

Utah requires disclosure in specified high-risk AI interactions and when a user asks directly whether they are interacting with AI.⁴² California and Washington have focused on establishing the provenance of synthetic media through embedded metadata, watermarks, and related detection tools.⁴³ California also requires developers to disclose information about the datasets used to train generative AI systems.⁴⁴

3.3.3 Frontier AI Safety

State regulation of frontier AI models has focused on transparency, reporting, and external oversight. California's SB 53 requires covered developers to publish safety frameworks and report serious safety incidents.⁴⁵ New York's Responsible AI Safety and Education Act concentrates on planning and reporting for catastrophic risks.⁴⁶ Illinois requires independent third-party safety audits.⁴⁷

41 S.B. 5, "An Act Concerning Online Safety" (Connecticut Artificial Intelligence Responsibility and Transparency Act), 2026 Reg. Sess. (Conn. 2026) (signed May 27, 2026; Pub. Act No. 26-15; effective dates staggered beginning Oct. 1, 2026).

42 Utah Code § 13-77-103(1)-(3), "Required Disclosures." <https://le.utah.gov/xcode/Title13/Chapter77/13-77-S103.html>.

43 California Business and Professions Code §§ 22757.2-22757.3, AI Transparency Act; Washington, Engrossed Second Substitute House Bill 1170, chap. 167, Laws of 2026, §§ 1-2, 8.

44 California Civil Code § 3111, "Artificial Intelligence Training Data Transparency."

45 S.B. 53, 2025-2026 Reg. Sess. (Cal. 2025) (signed Sept. 29, 2025; eff. Jan. 1, 2026); S.B. 149, "Artificial Intelligence Amendments," 2024 Gen. Sess. (Utah 2024) (enacted Mar. 13, 2024; effective May 1, 2024), <https://le.utah.gov/-2024/bills/static/SB0149.html>; S.B. 226, "Artificial Intelligence Consumer Protection Amendments," 2025 Gen. Sess. (Utah 2025) (signed Mar. 27, 2025; eff. May 7, 2025); S.B. 332, "Artificial Intelligence Policy Act Revisions," 2025 Gen. Sess. (Utah 2025) (signed Mar. 25, 2025; eff. May 7, 2025).

46 S. 6953-B, "Responsible AI Safety and Education Act" (RAISE Act), 2025-2026 Reg. Sess. (N.Y.) (signed Dec. 2025; chapter amendment Mar. 27, 2026; eff. Jan. 1, 2027), <https://www.nysenate.gov/legislation/bills/2025/S6953/amendment/B;S.8451-B/A.8962-B>, "New York Fundamental Artificial Intelligence Requirements in News Act" (NY FAIR News Act), 2025-2026 Reg. Sess. (N.Y.) (passed Senate June 8, 2026; passed Assembly June 8, 2026), <https://www.nysenate.gov/legislation/bills/2025/S8451/amendment/B>; Jennifer Johnson et al., "New York Legislature Passes Sweeping AI Safety Legislation," Global Policy Watch, June 24, 2025, <https://www.globalpolicywatch.com/2025/06/new-york-legislature-passes-sweeping-ai-safety-legislation>.

47 S.B. 315, 104th Gen. Assemb., Reg. Sess. (Ill. 2026) "Artificial Intelligence Safety Measures Act" (signed July 6, 2026; eff. Jan. 1, 2027).

4. ELECTION AND POLITICAL CONTENT

4.1 The First Amendment and AI-Generated Political Speech

In *United States v. Alvarez* (2012), the Supreme Court struck down a law criminalizing lying about military honors,⁴⁸ holding that deliberate falsehoods are generally protected speech unless the lie is an element of fraud, defamation, perjury, or a similarly established offense.⁴⁹

New AI capabilities are prompting a reassessment of false expression regulations. Shortly before the 2024 New Hampshire presidential primary, an AI-generated robocall impersonating President Biden asked voters to “save [their] vote for the November election” and claimed, “[y]our vote makes a difference in November, not this Tuesday.”⁵⁰ The Federal Communications Commission imposed a \$6 million fine on political consultant Steve Kramer, who commissioned the calls, and issued a ruling that calls using AI-generated or cloned voices are subject to the Telephone Consumer Protection Act’s restrictions on calls using an “artificial or prerecorded voice.”⁵¹

4.2 State Laws Governing Political Deepfakes

By June 2026, thirty-one states had passed laws that regulate deepfakes through disclosure requirements, prohibitions on specified deceptive content, or both.⁵²

48 *United States v. Alvarez*, 567 U.S. 709 (2012).

49 *Ibid.*

50 New Hampshire Department of Justice, “Steven Kramer Charged with Voter Suppression Over AI-Generated President Biden Robocalls,” May 23, 2024, <https://www.doj.nh.gov/news-and-media/steven-kramer-charged-voter-suppression-over-ai-generated-president-biden-robocalls>; Shannon Bond, “A Political Consultant Faces Charges and Fines for Biden Deepfake Robocalls,” NPR, May 23, 2024.

51 *In the Matter of Implications of Artificial Intelligence Technologies on Protecting Consumers from Unwanted Robocalls and Robotexts*, CG Docket No. 23-362, Declaratory Ruling, FCC 24-17 (rel. Feb. 8, 2024), <https://docs.fcc.gov/public/attachments/FCC-24-17A1.pdf>; FCC, “FCC Makes AI-Generated Voices in Robocalls Illegal,” February 8, 2024, <https://www.fcc.gov/document/fcc-makes-ai-generated-voices-robo-calls-illegal>.

52 “Tracker: State Legislation on Deepfakes in Elections,” Public Citizen, accessed June 30, 2026, <https://www.citizen.org/article/tracker-legislation-on-deepfakes-in-elections> (30 states have enacted laws as of May 2026, with Maryland the most recent enactment).

State deepfake laws predate the current wave of generative AI. Texas enacted the first election-specific law in 2019: SB 751 made it a Class A misdemeanor to create and cause the distribution of a deceptive deepfake video within thirty days of an election.⁵³ The law's scope is narrowed by its intent requirement—the video must be made to injure a candidate or influence an election—and by its timing window. No court has ruled on whether the provision is constitutional.⁵⁴ North Dakota's HB 1167 requires a disclaimer if political material uses AI to impersonate an individual.⁵⁵

Deepfake laws have been on the books for years, but litigation is only now beginning to define their constitutional limits. In August 2025, a federal district court held that California's AB 2839 imposed content, viewpoint, and speaker-based restrictions that failed strict scrutiny, and the court permanently enjoined enforcement of the law against the plaintiffs.⁵⁶ In January 2026, a court enjoined Hawaii's Act 191 as overbroad, finding that it swept in protected satire and parody.⁵⁷

53 S.B. 751, 86th Leg., R.S. (Tex. 2019), codified at Tex. Elec. Code Ann. § 255.004 (eff. Sept. 1, 2019).

54 S.B. 751, 86th Leg., R.S. (Tex. 2019), codified at Tex. Elec. Code Ann. § 255.004 (eff. Sept. 1, 2019).

55 H.B. 1167, "An Act to Create and Enact a New Section to Chapter 16.1-10 of the North Dakota Century Code, Relating to Artificial Intelligence Disclosure Statements," 69th Leg. Assemb. (N.D. 2025) (signed by Governor Apr. 11, 2025), <https://ndlegis.gov/assembly/69-2025/regular/bill-overview/bo1167.html>.

56 *Kohls v. Bonta*, 797 F. Supp. 3d 1177 - 2025.

57 *The Babylon Bee, LLC v. Lopez*, No. 1:25-cv-00234 (D. Haw. Jan. 30, 2026) (order granting summary judgment and permanent injunction); Delphine Strauss, "Hawaii's Deepfake Election Law Violates Free Speech, Court Finds," Bloomberg Law, January 30, 2026, <https://news.bloomberglaw.com/litigation/hawaiis-deepfake-election-law-violates-free-speech-court-finds>; see also Eugene Volokh, "Hawaii Deceptive Election-Related Deepfake Disclaimer Requirement Struck Down," Reason/Volokh Conspiracy, January 31, 2026, <https://reason.com/volokh/2026/01/31/hawaii-deceptive-election-related-deepfake-disclaimer-requirement-struck-down/>.

5. HATE SPEECH

Hate speech refers to offensive or derogatory words, signs, or symbols directed at a person or group based on identity characteristics like race, religion, ethnicity, gender, or sexual orientation. It often involves slurs or expressions of bigotry and can foment hostility and social discord.⁵⁸ The First Amendment protects hate speech broadly and has only a few, very narrow exceptions to this protection.⁵⁹ AI systems can generate such content at scale, leading some experts to ask whether existing doctrine is adequate and whether platform policies or state laws can fill the perceived constitutional gap.⁶⁰

AI developers and social media platforms, which are not bound by the First Amendment, can write, design, and enforce their own content moderation policies to filter out hate speech.⁶¹ Automated moderation is difficult to calibrate: overly broad filters can chill inquiry, suppress satire, and erase legitimate political commentary.

Laws in California and New York have sought to compel platforms to disclose their definitions and policies for moderating hate speech.⁶² Courts have found that such mandates may infringe on editorial discretion and amount to compelled speech.⁶³

In February 2025, California settled *X Corp. v. Bonta* by dropping the compelled disclosure

58 Anderson, Luvell and Michael Barnes, “Hate Speech”, The Stanford Encyclopedia of Philosophy (Summer 2025 Edition), Edward N. Zalta & Uri Nodelman (eds.) <https://plato.stanford.edu/archives/sum2025/entries/hate-speech/>.

59 *Brandenburg v. Ohio*, 395 U.S. 444 (1969); *Counterman v. Colorado*, 600 U.S. 66 (2023).

60 Jordi Calvet-Bademunt and Jacob Mchangama, “Freedom of Expression in Generative AI: A Snapshot of Content Policies,” The Future of Free Speech, February 2024, https://futurefreespeech.org/wp-content/uploads/2023/12/FFS_AI-Policies_Formatting.pdf.

61 Peter J. Benson and Valerie C. Brannon, “Section 230 Immunity and Generative Artificial Intelligence,” CRS Legal Sidebar LSB11097 (December 28, 2023), <https://www.congress.gov/crs-product/LSB11097>. Specifically, Section 230(c)(1) states that “[n]o provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.” 47 U.S. Code § 230, <https://www.law.cornell.edu/uscode/text/47/230>.

62 A.B. 587, “Social Media Companies: Terms of Service,” 2021-22 Reg. Sess. (Cal. 2022) (approved by Governor Sept. 13, 2022), https://leginfo.ca.gov/faces/billTextClient.xhtml?bill_id=20210220AB587.

63 *Moody v. NetChoice, LLC*, 603 U.S. (2024) (Kagan, J.) (holding that platforms’ content curation and moderation constitute First Amendment-protected editorial discretion, and that state laws compelling platforms to host or disclose content burden that discretion); see also *NetChoice, LLC v. Paxton*, consolidated with *Moody*, (remanding for further proceedings on facial challenges to Texas HB 20 and Florida SB 7072).

provisions of AB 587 in exchange for X’s voluntary transparency commitments.⁶⁴ While settlements do not make law, this one signaled to civil society and industry that compelling platforms to publish content-moderation policies raises serious First Amendment concerns.

In June 2025, X Corp. filed suit challenging New York’s law requiring social media platforms to disclose their hate speech and disinformation policies, arguing that it violates the First Amendment’s protection of editorial discretion. The outcome could have implications for AI content moderation transparency requirements.⁶⁵

64 *X Corp. v. Bonta*, No. 2:22-cv-01927-WBS-CKD (E.D. Cal. 2025): Final Judgment and Permanent Injunction (Feb. 24, 2025); *X Corp. v. Bonta*, 116 F.4th 888 (9th Cir. 2024) (preliminary injunction finding AB 587 content-category reporting provisions likely violate First Amendment right against compelled speech); Cal. Bus. & Prof. Code § 22677 (AB 587, 2022); Politico/Axios, “California agrees to drop parts of social media law challenged by Elon Musk’s X,” Feb. 24, 2025.

65 Nick Robins-Early, “Elon Musk’s X Sues New York over Hate Speech and Disinformation Law,” *The Guardian*, June 17, 2025, <https://www.theguardian.com/technology/2025/jun/17/elon-musk-new-york-hate-lawsuit-speech-law>.

6. DEFAMATION

6.1 Hallucinations and Developer Liability

In the United States, how the existing defamation framework applies to AI-generated material remains unsettled. AI systems frequently “hallucinate,” producing false statements and presenting them as fact.⁶⁶ A hallucination can become defamatory when the AI-generated false statement concerns a real person, raising the question whether developers bear liability comparable to that of traditional publishers.

Hallucinations were at issue in *Walters v. OpenAI*, brought by a Georgia radio host who alleged that ChatGPT falsely claimed he had embezzled funds from a nonprofit. The Superior Court of Gwinnett County, Georgia, granted summary judgment in favor of OpenAI.⁶⁷ The court found that “a reasonable reader” could not have concluded that the ChatGPT statements in question “communicated actual facts.”⁶⁸ The court also concluded that “Walters qualifies as a public figure”⁶⁹ and that he failed to demonstrate “knowing or reckless falsehood on OpenAI’s part.”⁷⁰ As of July 2026, no US court has awarded damages in a defamation case involving AI-generated speech.⁷¹

Individuals who are the subject of AI-generated false statements are not without recourse: settlement has become a common outcome. Settlements, however, prevent courts from ruling on the underlying liability or First Amendment questions. Political activist Robby

66 Chrystal R. China, “What Are AI Hallucinations?,” IBM, September 1, 2023, updated July 31, 2026, <https://www.ibm.com/think/topics/ai-hallucinations>.

67 Eugene Volokh, “OpenAI Wins Libel Lawsuit Brought by Gun Rights Activist Over Hallucinated Embezzlement Claims,” *Reason*, May 20, 2025, <https://reason.com/volokh/2025/05/20/openai-wins-libel-lawsuit-brought-by-gun-rights-activist-over-hallucinated-embezzlement-claims>.

68 *Walters v. OpenAI, L.L.C.*, No. 23-A-04860-2 (Ga. Super. Ct. Gwinnett County May 19, 2025) (order granting summary judgment), 6.

69 *Id.*, 13.

70 Volokh, “OpenAI Wins Libel Lawsuit,” *Id.*, 17.

71 Richard Epstein, “Suing OpenAI for ChatGPT-Produced Defamation Is a Futile Endeavor,” American Enterprise Institute, January 8, 2025, <https://www.aei.org/technology-and-innovation/suing-openai-for-chatgpt-produced-defamation-a-futile-endeavor/>; Eugene Volokh, “OpenAI Wins Libel Lawsuit Brought by Gun Rights Activist Over Hallucinated Embezzlement Claims,” *Reason*, May 20, 2025, <https://reason.com/volokh/2025/05/20/openai-wins-libel-lawsuit-brought-by-gun-rights-activist-over-hallucinated-embezzlement-claims>.

Starbuck, for example, sued Meta over allegedly defamatory statements its AI produced about him.⁷² The case was settled in August 2025, and Starbuck is now working with Meta to address “ideological and political bias” in its AI.⁷³ The arrangement is notable because a private settlement may shape the outputs of the system at issue.

6.2 The Applicability of Section 230 to AI-Generated Content

Section 230 of the Communications Decency Act, which has shaped the internet for decades, provides broad immunity to online platforms for content generated by third parties,⁷⁴ shielding them from defamation claims arising from user-generated content.

Efforts to narrow Section 230 have focused on online platforms, but there is also a legal debate over whether its protection extends to AI-generated content. The question for courts is whether a platform deploying generative AI tools becomes an “information content provider” of the resulting content. This question matters because, under the material contribution test, a provider loses immunity if it “materially contributes” to the development of unlawful content.⁷⁵

AI-generated content reflects design choices shaped by model fine-tuning, red teaming, policy guidelines, prompt engineering, and other technical decisions. The authors of Section 230 have explicitly stated that the provision would not shield AI chatbots.⁷⁶ Whether these design choices amount to editorial discretion the First Amendment protects, and whether that would place developers and deployers outside Section 230’s immunity, remains an open question.

72 Sarah Nassauer and Jacob Gershman, “Activist Robby Starbuck Sues Meta Over AI Answers About Him,” Wall Street Journal, April 29, 2025, <https://www.wsj.com/tech/ai/activist-robbie-starbuck-sues-meta-over-ai-answers-about-him-9eba5d8a>.

73 Joseph De Avila, “Meta, Robby Starbuck Settle AI Defamation Lawsuit,” Wall Street Journal, August 8, 2025, <https://www.wsj.com/tech/ai/meta-robbie-starbuck-ai-lawsuit-settlement-6c6e9b0a>.

74 Peter J. Benson and Valerie C. Brannon, “Section 230 Immunity and Generative Artificial Intelligence,” CRS Legal Sidebar LSB11097 (December 28, 2023), <https://www.congress.gov/crs-product/LSB11097>. Specifically, Section 230(c)(1) states that “[n]o provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.” 47 U.S. Code § 230, <https://www.law.cornell.edu/uscode/text/47/230>.

75 *Fair Housing Council v. Roommates.com, LLC*, 521 F.3d 1157, 1166, 1173–74 (9th Cir. 2008); *FTC v. Accusearch, Inc.*, 570 F.3d 1187, 1200 (10th Cir. 2008).

76 Cristiano Lima-Strong, “AI Chatbots Won’t Enjoy Tech’s Legal Shield, Section 230 Authors Say,” Washington Post, March 17, 2023, <https://www.washingtonpost.com/politics/2023/03/17/aichatbots-wont-enjoy-techs-legal-shield-section-230-authors-say>.

7. PROTECTION OF MINORS

US policymakers have focused considerable attention on interactions between AI systems and children, which is likely an extension of the scrutiny already directed at social media design features. That attention has taken three forms: product liability litigation over chatbot design; legislative efforts to mandate disclosure, age verification, and other restrictions on AI companions; and enforcement actions by state attorneys general.

7.1 AI Companion Liability and the Character.AI Litigation

Families have brought lawsuits against companion chatbot developers, alleging that prolonged chatbot use led to serious harm to children, including death by suicide.

Case Study: *Garcia v. Character Technologies: Chatbots and Product Liability*

Garcia v. Character Technologies is perhaps the most prominent product liability case involving AI.⁷⁷ In October 2024, the mother of a fourteen-year-old boy who died by suicide sued Character.AI, alleging that the product was negligently designed to maximize emotional dependency in vulnerable users without adequate safeguards.⁷⁸

Character.AI moved to dismiss on First Amendment grounds,⁷⁹ but the district court allowed the claims to proceed. The court found no identifiable “expressive choice” behind the chatbot’s responses—a test drawn from Justice Barrett’s concurrence in *Moody v. NetChoice*. At this stage of the litigation, Judge Anne Conway wrote, the court was “not prepared to hold that Character.AI’s output is speech.”⁸⁰ She did allow the company to

77 *Garcia v. Character Technologies*, 785 F. Supp. 3d 1157, 1178-1179 (M.D. Fla. 2025).

78 *Ibid.*

79 *Ibid.*

80 *Garcia v. Character Technologies*, 785 F. Supp. 3d 1157, 1178-1179 (M.D. Fla. 2025); Kate Payne, “In Lawsuit over Teen’s Death, Judge Rejects Arguments That AI Chatbots Have Free Speech Rights,” *AP News*, May 21, 2025, <https://apnews.com/article/ai-lawsuit-suicide-artificial-intelligence-free-speech-ccc77a5ff5a84bda753d2b044c83d4b6>.

assert “the First Amendment rights of its users” to receive the chatbot’s output.⁸¹ The ruling is an early procedural decision from a district court and binds no other jurisdiction. The *Garcia* case settled in early 2026,⁸² but the questions it raised remain. Treating chatbot output as a product rather than as speech opens a path to tort liability that the First Amendment would otherwise foreclose—a shift the Foundation for Individual Rights and Expression (FIRE) argued in an amicus brief carries “profound implications” for free speech.⁸³

While the case was pending, federal and state governments pursued their own inquiries into companion chatbots.⁸⁴ In August 2025, a bipartisan coalition of forty-four state attorneys general wrote to thirteen major AI companies, warning that they would use “all available legal and regulatory tools” to “ensure that children are protected from exploitation and harm.”⁸⁵ Soon thereafter, the FTC opened a market study under section 6(b) of the FTC Act, examining how seven AI companion companies monitor and protect minors.⁸⁶

In September 2025, a Senate subcommittee held a hearing on the harms of AI chatbots.⁸⁷ In November, Character.AI barred minors from open-ended conversations and strengthened

81 Reply in Support of Motion to Dismiss, *Garcia v. Character Technologies, Inc.*, No. 6:24-cv-01903-ACC-UAM (M.D. Fla. May 21, 2025), ECF No. 115, 27, <https://storage.courtlistener.com/recap/gov.uscourts.flmd.433581/gov.uscourts.flmd.433581.115.0.pdf>.

82 Character.AI, Google, and the individual defendants settled *Garcia v. Character Technologies* and four related cases in principle, Jan. 7, 2026 (terms confidential); CourtListener, *Garcia v. Character Technologies, Inc.*, No. 6:24-cv-01903-ACC-UAM, <https://www.courtlistener.com/docket/69300919/garcia-v-character-technologies-inc/> (jury trial scheduled Nov. 2, 2026); Open Class Actions, “Character.AI & Google Move to Settle Teen Chatbot Harm Lawsuits,” <https://openclassactions.com/lawsuits/character-ai-google-teen-harm-lawsuits-settlement.php> (accessed June 2026).

83 Foundation for Individual Rights and Expression, *Brief of Amicus Curiae, Garcia v. Character Technologies, Inc.*, No. 6:24-cv-01903-ACC-UAM (M.D. Fla. filed June 2025), <https://www.fire.org/cases/garcia-v-character-technologies-inc>; Foundation for Individual Rights and Expression, “FIRE to Court: AI Speech Is Still Speech — and the First Amendment Still Applies,” June 26, 2025, <https://www.thefire.org/news/fire-court-ai-speech-still-speech-and-first-amendment-still-applies>.

84 Natallie Rocha, “Google and Character.AI to Settle Lawsuit Over Teenager’s Death,” *The New York Times*, January 7, 2026, <https://www.nytimes.com/2026/01/07/technology/google-characterai-teenager-lawsuit.html>.

85 National Association of Attorneys General, “Bipartisan Coalition of State Attorneys General Issues Letter to AI Industry Leaders on Child Safety,” press release, August 25, 2025, <https://www.naag.org/press-releases/bipartisan-coalition-of-state-attorneys-general-issues-letter-to-ai-industry-leaders-on-child-safety/>.

86 Federal Trade Commission, “FTC Launches Inquiry into AI Chatbots Acting as Companions,” press release, September 11, 2025, <https://www.ftc.gov/news-events/news/press-releases/2025/09/ftc-launches-inquiry-ai-chatbots-acting-companions>; Federal Trade Commission, *6(b) Orders to File Special Report Regarding Advertising, Safety, and Data Handling Practices by Companies Offering Generative Artificial Intelligence (“AI”) Companion Products or Services*, Sept. 2025, <https://www.ftc.gov/reports/6b-orders-file-special-report-regarding-advertising-safety-data-handling-practices-companies>.

87 *Examining the Harm of AI Chatbots: Hearing Before the Subcommittee on Crime and Counterterrorism of the Senate Committee on the Judiciary*, 119th Cong. (Sept. 16, 2025) (S. Hrg. 119-256), <https://www.congress.gov/event/119th-congress/senate-event/LC75088/text>; Testimony of Megan Garcia, id., <https://www.judiciary.senate.gov/imo/media/doc/e2e8fc50-a9ac-05ec-edd7-277cb0afcdf2/2025-09-16%20PM%20-%20Testimony%20-%20Garcia.pdf>.

its age assurance controls.⁸⁸ Regulatory pressure had again produced a policy change without a law being passed.

Texas had opened an investigation into Character.AI in December 2024.⁸⁹ The day after the *Garcia* settlement, Kentucky went further, becoming the first state to sue the company over alleged deceptive trade practices and predatory design.⁹⁰

The product liability theory at the core of *Garcia* has spread beyond Character.AI. In June of 2026, the state filed a civil suit against the company over the safety of its products.⁹¹ Separately, in April 2026, Florida's attorney general launched a criminal investigation into OpenAI over ChatGPT's alleged role in a mass shooting at a Florida university.⁹²

7.2 Federal Legislative Response: the KIDS Act

As of mid-2026, Congress has enacted no child safety legislation, despite an active session.⁹³

The House of Representatives passed the Kids Internet and Digital Safety (KIDS) Act on June 29, 2026: a package bundling numerous child online safety provisions.⁹⁴ It would have set a federal floor, using a conflict preemption standard that would have displaced state law only where the two could not both be followed—preserving state laws that provided stronger

88 Character.AI, "Important Changes for Teens on Character.ai," updated October 29 2025, <https://support.character.ai/hc/en-us/articles/42645561782555-Important-Changes-for-Teens-on-Character-ai>.

89 Office of Texas Attorney General Ken Paxton, "Attorney General Ken Paxton Investigates Meta and Character.AI for Misleading Children with Deceptive AI-Generated Mental Health Services," press release, August 18, 2025, <https://www.texasattorneygeneral.gov/news/releases/attorney-general-ken-paxton-investigates-meta-and-characterai-misleading-children-deceptive-ai>.

90 Office of Kentucky Attorney General, "AG Coleman Sues AI Chatbot Company for Preying on Children," January 8, 2026, <https://www.kentucky.gov/Pages/Activity-stream.aspx?n=AttorneyGeneral&prld=1857>; *Commonwealth of Ky. ex rel. Russell Coleman, Att'y Gen. v. Character Techs., Inc.*, No. 26-CI-00029 (Ky. Cir. Ct. Jan. 8, 2026).

91 Hadas Gold, "Florida Sues OpenAI, Alleging It's Unsafe for Children," CNN, June 2, 2026, <https://www.cnn.com/2026/06/01/business/florida-sues-chatgpt-openai-sam-altman>.

92 Office of the Florida Attorney General, "Attorney General James Uthmeier Launches Criminal Investigation into OpenAI and ChatGPT," April 21, 2026, <https://www.myfloridalegal.com/newsrelease/attorney-general-james-uthmeier-launches-criminal-investigation-openai-chatgpt>; Jared Perlo and Abigail Brooks "Florida's Attorney General Announces Criminal Investigation into OpenAI," NBC News, April 21, 2026, <https://www.nbcnews.com/tech/tech-news/florida-attorney-general-criminal-investigation-openai-fsu-chatgpt-rcna341205>.

93 Emma Whitford, "US House Child Safety Plan Shows Movement on Preemption, 'Knowledge,'" MLex, June 23, 2026, <https://www.mlex.com/mlex/articles/2493019>; House Energy and Commerce Committee, Kids Internet and Digital Safety Act (KIDS Act), discussion draft published June 23, 2026 (including KOSA, SAFE BOTs Act, COPPA 2.0, and SCREEN Act).

94 H.R. 7757, Kids Internet and Digital Safety (KIDS) Act, 119th Cong. (passed House June 29, 2026, 267-117), <https://www.congress.gov/bill/119th-congress/house-bill/7757>; House Energy and Commerce Committee, press release, June 22, 2026, The bill includes an updated Kids Online Safety Act, the Safeguarding Adolescents From Exploitative BOTs Act (requiring platforms to disclose that chatbots are not human), COPPA 2.0, and the SCREEN Act (requiring age verification for platforms where at least one-third of content is sexual material harmful to minors).

child protections. Senate sponsors declared the bill dead on arrival over its omission of a duty of care provision (a legal requirement that platforms take reasonable steps to prevent harms).⁹⁵

The KIDS Act would have regulated speech in several ways that raise free expression concerns. The Center for Democracy and Technology (CDT) warned that the age verification requirements would have pressured online service providers to verify the age of all users, causing significant privacy consequences.⁹⁶ Moreover, the Electronic Frontier Foundation (EFF) cautioned that the bill would have pressured “companies to police lawful speech online.”⁹⁷

7.3 State Chatbot Laws and Design Restrictions

States have moved faster than the federal government to regulate chatbots. As of mid-2026, fifteen have enacted companion-chatbot or chatbot-safety laws governing how those systems interact with minors, and many more bills remain pending. The legislative effort has drawn broad bipartisan support.⁹⁸

Two state laws are worth closer attention: California’s SB 243 and Georgia’s SB 540. California’s law, like many of its technology regulations, has served as a template for other states. It requires platforms to notify known minor users every three hours that they are interacting with a chatbot, to implement self-harm prevention protocols, and to refer users expressing suicidal ideation to crisis services.⁹⁹ Georgia’s Conversational Artificial Intelligence Safety Act, effective July 1, 2027, requires disclosure at the start of chatbot sessions, prohibits chatbots from simulating romantic relationships with minors or using manipulative engagement techniques, and mandates crisis protocols for conversations

95 Allison Mollenkamp, “Kids Bill Faces Uncertainty After House Passage,” *Roll Call*, June 29, 2026, <https://rollcall.com/2026/06/29/kids-bill-faces-uncertainty-after-house-passage/>; see also Crowell and Moring, “House Passes KIDS Act H.R. 7757: What Online Platforms Must Know Before Senate Showdown,” June 30, 2026, <https://www.crowell.com/en/insights/client-alerts/house-advances-bipartisan-kids-online-safety-bill-but-senate-showdown-looms>.

96 Center for Democracy & Technology, Letter to Senate Commerce Committee, August 4, 2026, https://cdt.org/wp-content/uploads/2026/08/Commerce-Letter_Updated.pdf.

97 India McKinney, “The House Passed The KIDS Act—The Senate Should Reject It,” Electronic Frontier Foundation, July 9, 2026, <https://www.eff.org/deeplinks/2026/07/house-passed-kids-act-senate-should-reject-it>.

98 Future of Privacy Forum, “2026 Chatbot Legislation Tracker,” <https://fpf.org/2026-chatbot-legislation-tracker/>.

99 Cal. S.B. 243 (signed Oct. 13, 2025; eff. Jan. 1, 2026).

involving suicidal ideation or self-harm.¹⁰⁰

A growing number of states have enacted or are advancing age verification requirements for online services and AI platforms.¹⁰¹ The Supreme Court's decision in *Free Speech Coalition v. Paxton* (2025) upheld Texas's age verification law for websites where more than one-third of the content is "harmful to minors."¹⁰² States have since introduced or enacted age verification bills extending to AI services and social media platforms accessible to minors. Several have faced First Amendment challenges.¹⁰³

These laws pose particular risks for LGBTQ+ expression. Civil liberties groups, including the American Civil Liberties Union (ACLU) and EFF, warn that state age verification laws burden adults' access to lawful speech and jeopardize their anonymity and data security online.¹⁰⁴ Moreover, LGBT Tech, a nonpartisan coalition working on technology policy for LGBT communities, has argued that the vague "harmful to minors" standard could be turned against LGBTQ+ content.¹⁰⁵ In states with definitions broad enough to encompass sexual health information and LGBTQ+ educational material, age verification requirements risk extending well beyond the explicit content they nominally target.¹⁰⁶

These laws have already had far-reaching consequences. After Texas's age verification law took effect in 2024, several adult content platforms blocked access in the state entirely, and the same pattern has followed elsewhere as comparable laws have taken effect.¹⁰⁷

100 Ga. S.B. 540, Conversational Artificial Intelligence Safety Act (signed May 11, 2026; eff. July 1, 2027) (chatbot disclosure, minor safety restrictions, crisis-response protocols; no carve-out for embedded platform chatbots); LegalTech Digest, "Georgia Enacts Sweeping AI Companion Law Without Big-Tech Exemptions," May 21, 2026.

101 *Free Speech Coalition, Inc. v. Paxton*, 606 U.S. 461 (2025); National Conference of State Legislatures, "Artificial Intelligence 2025 Legislation," accessed June 30, 2026, <https://www.ncsl.org/technology-and-communication/artificial-intelligence-2025-legislation>.

102 Ibid.

103 Peter J. Benson and Valerie C. Brannon, *First Amendment Challenges to Regulation of Social Media*, Congressional Research Service, R49308 (August 24, 2026), <https://www.congress.gov/crs-product/R49308>; *NetChoice v. Hilgers*, No. 4:26-CV-3149 (D. Neb. June 27, 2026).

104 American Civil Liberties Union, "ACLU Comment on Supreme Court Decision in *Free Speech Coalition v. Paxton*," June 27, 2025, <https://www.aclu.org/press-releases/fsc-paxton-age-verification>; Electronic Frontier Foundation, "The Supreme Court's Decision on Age Verification Tramples Free Speech and Undermines Privacy," December 10, 2025, <https://www.eff.org/pages/supreme-courts-decision-age-verification-tramples-free-speech-and-undermines-privacy>.

105 LGBT Tech, "The Supreme Court's *Free Speech Coalition v. Paxton* Decision Is a Privacy Nightmare for LGBTQ+ People," June 27, 2025, <https://www.lgbtttech.org/post/the-supreme-court-s-free-speech-coalition-v-paxton-decision-is-a-privacy-nightmare-for-lgbtq-peopl>.

106 Ibid.

107 Christopher Kane, "Pornhub blocks Texas accessing site over age verification law," *Washington Blade*, March 15, 2024, <https://www.washingtonblade.com/2024/03/15/pornhub-blocks-texas-accessing-site-over-age-verification-law/>.

8. EXPLICIT CONTENT

Child sexual abuse material (CSAM) and nonconsensual intimate imagery (NCII) are the subject of broad legal and political consensus, and both federal and state law regulate them. The TAKE IT DOWN Act, enacted in May 2025, is the principal federal measure, and nearly every state has passed legislation of its own.

8.1 The TAKE IT DOWN Act

The TAKE IT DOWN Act, passed nearly unanimously by Congress and signed into law by President Trump in May 2025, prohibits the distribution of AI-generated CSAM and NCII.¹⁰⁸ Its coverage extends beyond “sexually explicit” images to nude images published with the intent to “abuse, humiliate, harass, or degrade” a minor.¹⁰⁹ The law requires large online platforms to establish a process for victims to report such distribution and to remove AI-generated NCII within forty-eight hours of a request.¹¹⁰ Those platform compliance requirements took effect in May 2026.

The TAKE IT DOWN Act is well intentioned and addresses serious harms, especially to children, women, and LGBTQ+ individuals. Yet caution is warranted. Civil liberties groups warn that without narrowly tailored exemptions, the law could suppress speech that is constitutionally protected and falls outside its intended scope.¹¹¹ They also note that the penalty structure incentivizes platforms to over-remove content rather than risk liability.¹¹²

The Future of Free Speech has noted that the law’s framing obscures its capacity to serve

108 Tools to Address Known Exploitation by Immobilizing Technological Deepfakes on Websites and Networks (TAKE IT DOWN) Act, Pub. L. No. 119-12, 139 Stat. 55 (2025), <https://www.govinfo.gov/app/details/PLAW-119publ12>.

109 Sunny Gandhi and Adam Billen, “The US Senate’s Passage of the TAKE IT DOWN ACT Is Progress on an Urgent, Growing Problem,” *Tech Policy Press*, February 21, 2025, <https://www.techpolicy.press/the-us-senates-passage-of-the-take-it-down-act-is-progress-on-an-urgent-growing-problem/>.

110 Tools to Address Known Exploitation by Immobilizing Technological Deepfakes on Websites and Networks (TAKE IT DOWN) Act, Pub. L. No. 119-12, 139 Stat. 55 (2025), <https://www.govinfo.gov/app/details/PLAW-119publ12>.

111 Center for Democracy and Technology et al., “Letter Expressing Concerns Regarding the TAKE IT DOWN Act,” CDT, February 12, 2025, https://cdt.org/wp-content/uploads/2025/02/TAKE-IT-DOWN-Sign-On-Letter_21225.pdf.

112 Ibid.

as a censorship instrument,¹¹³ a risk that runs through its enforcement authority. The Act “grants the FTC broad discretion to investigate, penalize, and potentially litigate against online platforms.”¹¹⁴ Given that the administration has repeatedly threatened legal action against critics and media outlets on ideological grounds,¹¹⁵ the Act carries a structural risk of suppressing lawful political satire, investigative journalism, and LGBTQ+ content.¹¹⁶

8.2 AI-Generated Child Sexual Abuse Material

In the United States, CSAM involving real minors is unprotected speech subject to criminal prohibition, a point on which there is strong legal consensus. Federal law has adapted to new technologies: the PROTECT Act (2003) extended it to cover virtual CSAM¹¹⁷ after the Supreme Court struck down an earlier, overbroad statute in *Ashcroft v. Free Speech Coalition*.¹¹⁸

The legal status of AI-generated or computer-edited CSAM that does not depict actual children is more complicated. In *United States v. Anderegg* (2026), a federal appeals court held that private possession of AI-generated CSAM not depicting a real child was protected under *Stanley* and *Ashcroft*.¹¹⁹ Judge Lee wrote, “Given the relentless advancement in artificial intelligence models, we have some concerns about the lines these cases draw, but we are not free to redraw them ourselves.”¹²⁰ Any change would have to come from the Supreme Court. The TAKE IT DOWN Act, which targets distribution, does not prohibit personal possession.

In addition to federal regulation, nearly every state has enacted a law addressing AI-

113 Ashkhen Kazaryan and Ashley Haek, “The Road to Enforcement Chaos: The Hidden Dangers of the TAKE IT DOWN Act,” *The Bedrock Principle / Future of Free Speech*, May 12, 2025, <https://www.bedrockprinciple.com/p/the-road-to-enforcement-chaos-the>.

114 Ibid.

115 Wikipedia, “Donald Trump’s Conflict with the News Media,” under “Retaliatory Lawsuits,” accessed August 24, 2026, https://en.wikipedia.org/wiki/Donald_Trump%27s_conflict_with_the_news_media#Retaliatory_lawsuits.

116 Ashkhen Kazaryan and Ashley Haek, “The Road to Enforcement Chaos: The Hidden Dangers of the TAKE IT DOWN Act,” *The Bedrock Principle / Future of Free Speech*, May 12, 2025, <https://www.bedrockprinciple.com/p/the-road-to-enforcement-chaos-the>.

117 Prosecutorial Remedies and Other Tools to End the Exploitation of Children Today (PROTECT) Act of 2003, Pub. L. No. 108-21, 117 Stat. 650 (Apr. 30, 2003), <https://www.govinfo.gov/app/details/PLAW-108publ21>.

118 *New York v. Ferber*, 458 U.S. 747 (1982); *Ashcroft v. Free Speech Coalition*, 535 U.S. 234 (2002); PROTECT Act of 2003, <https://www.govinfo.gov/app/details/PLAW-108publ21>.

119 Nitasha Tiku, “Federal Judge Warns Law Is Being Left Behind by AI Child Sex Abuse Images,” *The Washington Post*, August 26, 2026, <https://www.washingtonpost.com/technology/2026/08/26/federal-judge-warns-law-is-being-left-behind-by-ai-sex-abuse-images/>; *United States v. Anderegg*, No. 25-1354 (7th Cir. Aug. 25, 2026); *Stanley v. Georgia*, 394 U.S. 557, 564 (1969); and *Ashcroft v. Free Speech Coalition*, 535 U.S. 234 (2002).

120 *Anderegg*, No. 25-1354, slip op. at 3.

generated CSAM.¹²¹ California’s and Illinois’s statutes include AI-generated content in their definitions of CSAM. Others have enacted laws targeting AI-generated abuse imagery—Montana’s HB 82, for instance, criminalizes the production, distribution, and possession of computer-generated CSAM.¹²²

8.3 AI-Generated Non-Consensual Intimate Imagery

The TAKE IT DOWN Act is the leading federal response to deepfake pornography and AI-generated intimate forgeries.

Before and after its enactment, states passed legislation covering NCII, including AI-generated intimate deepfakes.¹²³ California, New York, Virginia, Texas, and Minnesota provide civil or criminal remedies, or both, for the unauthorized distribution of synthetic sexually explicit images. Colorado’s SB 25-288 takes a different approach, creating a private right of action for nonconsensual disclosure of, or threat to disclose, an intimate digital depiction created by generative AI, image editing software, or computer-generated means.¹²⁴

¹²¹ “State Laws Criminalizing AI-Generated or Computer-Edited CSAM,” Enough Abuse, n.d., accessed June 2026, <https://enoughabuse.org/get-vocal/laws-by-state/state-laws-criminalizing-ai-generated-or-computer-edited-child-sexual-abuse-material-csam> (46 states as of August 31 2026).

¹²² H.B. 82, “An Act Generally Revising Crimes Against Children; Creating the Offense of Grooming of a Child for a Sexual Offense,” 69th Leg. (Mont. 2025) (signed by Governor Apr. 7, 2025; effective July 1, 2025), <https://bills.legmt.gov/#/laws/bill/2/LC0232>.

¹²³ “Tracker: State Legislation on Intimate Deepfakes,” Public Citizen, accessed June 30 2026, <https://www.citizen.org/article/tracker-intimate-deepfakes-state-legislation/>.

¹²⁴ S.B. 288, “Intimate Digital Depictions Criminal & Civil Actions,” 75th Gen. Assemb. (Colo. 2025) (signed by Governor June 2, 2025), <https://leg.colorado.gov/bills/sb25-288>.

9. COPYRIGHT

The United States is still determining how existing copyright law applies to AI-generated work. Two questions remain open. First, does training models on copyrighted works qualify as fair use? Second, do AI outputs that replicate training data constitute infringement?

9.1 Training Data, Fair Use, and the Anthropic Settlement

Litigation has centered on whether training models on copyrighted works qualifies as fair use, with divergent results across a rapidly growing docket.

In an early case, *Thomson Reuters v. Ross Intelligence*, the District Court for the District of Delaware rejected a fair use defense involving the use of copyrighted legal headnotes to train a non-generative legal research tool.¹²⁵ In *Bartz v. Anthropic*, by contrast, Judge William Alsup ruled that training Anthropic’s Claude model on copyrighted books qualified as fair use.¹²⁶ Judge Alsup reasoned that the use was “quintessentially transformative”¹²⁷ because it enabled the generation of new text.¹²⁸ The court nonetheless found Anthropic potentially liable for storing more than seven million pirated books in a centralized library. *Bartz* thus distinguishes between use and acquisition: training on books is fair use, but obtaining them by piracy is not.

In September 2025, Anthropic reached a \$1.5 billion settlement with the author plaintiffs, the largest copyright class action settlement to date. It covers almost 500,000 titles downloaded from pirate sources, at about \$3,000 per title. How AI companies acquire

125 *Thomson Reuters Enterprise Centre GmbH v. ROSS Intelligence Inc.*, No. 1:20-cv-00613-SB (D. Del. Feb. 11, 2025), https://www.ded.uscourts.gov/sites/ded/files/opinions/20-613_5.pdf; Monique N. Bhargava, Katherine Litaker, “Court Shuts Down AI Fair Use Argument in Thomson Reuters Enterprise Centre GMBH v. Ross Intelligence Inc.,” *Reed Smith*, March 3, 2025, <https://www.reedsmith.com/en/perspectives/2025/03/court-ai-fair-use-thomson-reuters-enterprise-gmbh-ross-intelligence>.

126 *Bartz et al. v. Anthropic PBC*, No. 3:24-cv-05417-WHA (N.D. Cal.) (preliminary approval Sept. 25, 2025; final approval July 20, 2026).

127 *Bartz et al. v. Anthropic PBC*; Authors Guild, “Bartz v. Anthropic Settlement: What Authors Need to Know,” <https://authorsguild.org/advocacy/artificial-intelligence/what-authors-need-to-know-about-the-anthropic-settlement/> (accessed June 2026).

128 Andrew Jeong, “Federal Court Says Copyrighted Books Are Fair Use for AI Training,” *The Washington Post*, June 25, 2025, <https://www.washingtonpost.com/technology/2025/06/25/ai-copyright-anthropic-books>.

training data may therefore matter as much as what they do with it: lawfully acquired training data may qualify as fair use, whereas pirated libraries may not.¹²⁹

Two other suits have not produced a fair use ruling. A district court dismissed a lawsuit against Meta after finding that the plaintiffs had failed to articulate a viable legal theory, without determining whether Meta's conduct was lawful.¹³⁰ *The New York Times Co. v. Microsoft Corp.* lawsuit against OpenAI remains pending, and in early 2026 the court ordered OpenAI to produce tens of millions of output logs in discovery.¹³¹

Outside the courts, the US Copyright Office addressed generative AI training and fair use in the third installment of its May 2025 report *Copyright and Artificial Intelligence*. It concludes that “training a generative AI foundation model on a large and diverse dataset will often be transformative,” and recommends allowing the licensing market to develop without government intervention.¹³² The Trump administration fired the head of the Copyright Office the day after the report's release, with Congressional Democrats attributing the move to her refusal to “rubber-stamp” efforts to mine copyrighted works for AI training.¹³³

9.2 Copyright Eligibility and Human Authorship

Copyright law requires human authorship for eligibility.¹³⁴ In *Thaler v. Perlmutter*, the US District Court for the District of Columbia upheld the Copyright Office's refusal to register

129 *Bartz et al. v. Anthropic PBC*, No. 3:24-cv-05417-WHA (N.D. Cal.) (preliminary approval Sept. 25, 2025; final approval May 14, 2026; \$1.5 billion settlement); Authors Guild, “Bartz v. Anthropic Settlement: What Authors Need to Know,” <https://authorsguild.org/advocacy/artificial-intelligence/what-authors-need-to-know-about-the-anthropic-settlement/> (accessed June 30, 2026).

130 Dan Milmo, “Meta Wins AI Copyright Lawsuit as US Judge Rules Against Authors,” *The Guardian*, June 26, 2025, <https://www.theguardian.com/technology/2025/jun/26/meta-wins-ai-copyright-lawsuit-as-us-judge-rules-against-authors>; Kate Knibbs, “A Judge Says Meta's AI Copyright Case Is About ‘the Next Taylor Swift,’” *Wired*, May 1, 2025, <https://www.wired.com/story/meta-lawsuit-copyright-hearing-artificial-intelligence>.

131 *In re OpenAI, Inc. Copyright Infringement Litigation* (S.D.N.Y.) (order denying motion to dismiss Oct. 27, 2025; orders compelling production of output logs Jan. 5 and Mar. 9, 2026); *New York Times Co. v. Microsoft Corp. et al.*, No. 1:23-cv-11195 (S.D.N.Y.) (ongoing as of June 30, 2026).

132 US Copyright Office, “Copyright and Artificial Intelligence, Part 3: Generative AI Training (Pre-Publication Version),” in Report on Copyright and Artificial Intelligence (Washington, DC: US Copyright Office, May 2025), <https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-3-Generative-AI-Training-Report-Pre-Publication-Version.pdf>.

133 Matt O'Brien, “Trump Administration Fires Top Copyright Official Days after Firing Librarian of Congress,” *AP News*, May 11, 2025, <https://apnews.com/article/copyright-director-firing-government-trump-7ab99992a96131bce7de853b66feec68>.

134 US Copyright Office, “Copyright and Artificial Intelligence, Part 2: Copyrightability,” in Report on Copyright and Artificial Intelligence, (Washington, DC: US Copyright Office, January 2025), <https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-2-Copyrightability-Report.pdf>.

a visual artwork that was generated solely by an AI system.¹³⁵ The Supreme Court later declined to hear an appeal, leaving the ruling in place.¹³⁶

Part 2 of the Copyright Office’s *Copyright and Artificial Intelligence* report affirms the necessity of substantial human involvement. The Office concludes that a work may qualify for copyright protection where a human meaningfully selects, arranges, or modifies AI-generated elements, but that content generated autonomously by an AI system is not copyrightable. Prompts alone, the report states, do not confer authorship over AI output.¹³⁷

9.3 AI Output Infringement

A second question concerns AI outputs and infringement. In *New York Times Co. v. Microsoft Corp.*, the plaintiffs allege that outputs reproducing passages verbatim from training data constitute direct infringement.¹³⁸

9.4 Digital Likeness Protection

A person’s likeness falls outside traditional copyright protections, so some legislatures have instead protected it directly by targeting unauthorized commercial use.

Tennessee’s General Assembly led with the Ensuring Likeness, Voice, and Image Security (ELVIS) Act. Other states have since enacted similar laws restricting the nonconsensual AI-generated use of a person’s likeness in commercial or misleading contexts.

At the federal level, the NO FAKES Act was introduced in 2025 but has not passed.¹³⁹ It has drawn significant opposition from civil liberties groups. In June 2026, CDT led a coalition urging the Senate Judiciary Committee not to advance the bill in its current form; signatories included The Future of Free Speech, EFF, FIRE, Public Knowledge, and the

¹³⁵ *Thaler v. Perlmutter*, 687 F. Supp. 3d 140, 142 (D.D.C. 2023).

¹³⁶ *Thaler v. Perlmutter*, 130 F.4th 1039 (D.C. Cir. 2025).

¹³⁷ US Copyright Office, “Copyright and Artificial Intelligence, Part 2: Copyrightability,” in Report on Copyright and Artificial Intelligence, (Washington, DC: US Copyright Office, January 2025), <https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-2-Copyrightability-Report.pdf>.

¹³⁸ *New York Times Co. v. Microsoft Corp. et al.*, No. 1:23-cv-11195 (S.D.N.Y.) (ongoing as of June 30, 2026). The case remains in active discovery and is expected to proceed toward trial in late 2026 or 2027.

¹³⁹ Nurture Originals, Foster Art, and Keep Entertainment Safe Act of 2025, S. 1367, 119th Cong., 1st sess. (2025), <https://www.congress.gov/bills/119th/congress/senate-bill/1367>.

ARTIFICIAL INTELLIGENCE AND FREEDOM OF EXPRESSION IN THE UNITED STATES

Organization for Transformative Works.¹⁴⁰ The coalition warned that the bill's steep penalties for hosting contested content would cause platforms to over-remove lawful expression, including satire, parody, commentary, and news.¹⁴¹

¹⁴⁰ Center for Democracy and Technology et al., Letter to Senator Chuck Grassley and Senate Judiciary Committee in Opposition to S. 4591, the NO FAKES Act of 2026, June 16, 2026, https://cdt.org/wp-content/uploads/2026/06/NO-FAKES-Summer-2026_FINAL-1.pdf; Center for Democracy and Technology, "CDT Leads Coalition Opposing the NO FAKES Act," June 16, 2026, <https://cdt.org/insights/cdt-leads-coalition-opposing-the-no-fakes-act/>.

¹⁴¹ Ibid.

10. AI-POWERED SURVEILLANCE

This section examines government use of AI for surveillance and its implications for free expression.

10.1 DOGE and Federal AI Surveillance: Political Loyalty Monitoring

Federal and local governments have significantly expanded their use of AI-powered surveillance tools, with documented chilling effects on expression and assembly. The Department of Government Efficiency (DOGE), established by executive order in January 2025, reportedly used AI to surveil federal agency communications for perceived hostility to the Trump administration and ordered staff to identify employees not “loyal” to its agenda.¹⁴² Grok, the chatbot developed by Elon Musk’s xAI and integrated into X (formerly Twitter), has also been deployed across federal agencies. Because Musk served as a special government employee in 2025, the deployment has raised conflict-of-interest and data security concerns.¹⁴³

10.2 ICE, Facial Recognition, and the Chilling of Expression

The Department of Homeland Security (DHS), including Immigration and Customs Enforcement (ICE), has expanded its use of AI surveillance.¹⁴⁴ Mobile Fortify, a real-time facial recognition tool, has been deployed during protests and enforcement operations where bystanders and journalists were present.¹⁴⁵ Homeland Security Secretary Markwayne

¹⁴² Alexandra Ulmer et al., “Musk’s DOGE Using AI to Snoop on U.S. federal workers, sources say,” *Reuters*, April 8, 2025, <https://www.reuters.com/technology/artificial-intelligence/musks-doge-using-ai-snoop-us-federal-workers-sources-say-2025-04-08/>; American Civil Liberties Union, “ACLU Seeks Records on DOGE’s Unrestricted Access to Americans’ Data,” February 7, 2025, <https://www.aclu.org/press-releases/aclu-seeks-records-on-doges-unrestricted-access-to-americans-data-urges-congress-to-step-up>.

¹⁴³ Marisa Taylor and Alexandra Ulmer, “Musk’s DOGE Expanding his Grok AI in US Government, Raising Conflict Concerns,” *Reuters*, May 23, 2025, <https://www.reuters.com/sustainability/boards-policy-regulation/musks-doge-expanding-his-grok-ai-us-government-raising-conflict-concerns-2025-05-23/>.

¹⁴⁴ Lindsey Wilkinson and Madison Alder, “ICE Drives AI Use Case Growth within Homeland Security,” *FedScoop*, January 28, 2026, <https://fedscoop.com/dhs-ai-inventory-mobile-fortify-palantir/>.

¹⁴⁵ Anthony Kimery, “Lawsuit Casts New Light on ICE, CBP’s Expanding Biometric, Visual Surveillance Dragnet,” *Biometric Update*, January 13, 2026, <https://www.biometricupdate.com/202601/lawsuit-casts-new-light-on-ice-cbps-expanding-biometric-visual-surveillance-dragnet>.

Mullin confirmed in a congressional hearing that DHS has used facial recognition to identify people who attended protests.¹⁴⁶

10.3 Automated License Plate Reader Networks

More than 5,000 law enforcement agencies use Flock Safety’s automated license plate recognition (ALPR) cameras, and the network conducts more than 20 billion vehicle scans monthly.¹⁴⁷ Local police have conducted searches on behalf of federal immigration authorities.¹⁴⁸ In response, Washington state passed SB 6002, which limits the types of agencies that can use ALPRs and bars them from using cameras for immigration enforcement or to surveil protesters, facilities providing protected health care, or other protected activities.¹⁴⁹

10.4 The Regulatory Gap

No comprehensive federal statute limits AI-powered surveillance by government agencies.¹⁵⁰ The Privacy Act of 1974 regulates how federal agencies maintain and disclose records about individuals. It also restricts agencies from maintaining records that describe how individuals exercise First Amendment rights.¹⁵¹

The Fourth Amendment also imposes some constraints, but whether and how it applies to AI-assisted surveillance remains unsettled. In June 2026, the Supreme Court expanded its protections in *Chatrue v. United States*,¹⁵² holding that a person retains a reasonable expectation of privacy in location history data and that law enforcement conducts a Fourth

146 Meg Anderson, “Some Local Police Have Access to an ICE Facial Recognition App”, *NPR*, June 19, 2026, republished by Delaware Public Media, <https://www.delawarepublic.org/npr-headlines/2026-06-19/some-local-police-have-access-to-an-ice-facial-recognition-app>.

147 AI Landes, License Plate Cameras Scan 20 Billion Vehicles Monthly—Cities Are Canceling Contracts, *Gadget Review*, June 30, 2026, <https://www.gadgetreview.com/license-plate-cameras-scan-20-billion-vehicles-monthly-cities-are-canceling-contracts>.

148 Jason Koebler and Joseph Cox, ICE Taps into Nationwide AI-Enabled Camera Network, Data Shows, *404 Media*, May 27, 2025, <https://www.404media.co/ice-taps-into-nationwide-ai-enabled-camera-network-data-shows/>.

149 Washington, Engrossed Substitute Senate Bill 6002, chap. 239, Laws of 2026.

150 5 U.S.C. § 552a (Privacy Act of 1974); Jay Stanley, “Machine Surveillance Is Being Super-Charged by Large AI Models,” *American Civil Liberties Union*, March 21, 2025, <https://www.aclu.org/news/privacy-technology/machine-surveillance-is-being-super-charged-by-large-ai-models>.

151 Zack Whittaker, “FBI Is Buying Location Data to Track US Citizens, Director Confirms,” *TechCrunch*, March 18, 2026, <https://techcrunch.com/2026/03/18/fbi-is-buying-location-data-to-track-us-citizens-kash-patel-wyden/>; Lindsey Wilkinson, “Privacy Advocates Sound Alarm on ‘Data Broker Loophole’ Used by FBI, Other Federal Agencies,” *FedScoop*, March 20, 2026, <https://fedscoop.com/fbi-data-broker-loop-hole-purchase-dhs/>; *Worldwide Threats: Hearing Before the Senate Select Committee on Intelligence*, 119th Cong., 2d sess. (March 18, 2026) (testimony of FBI Director Kash Patel).

152 *Chatrue v. United States*, 609 U.S. ____ (2026).

Amendment search when it obtains that data through a geofence warrant.¹⁵³ Geofence warrants can directly affect First Amendment rights. A warrant that gathers information on who attended a protest, a house of worship, a reproductive health clinic, or another site of expressive association produces what is functionally a membership list. Under *Zurcher v. Stanford Daily*, the Fourth Amendment’s probable cause and particularity requirements must be applied with “scrupulous exactitude” in such circumstances.¹⁵⁴

10.5 The Surveillance Threat to Free Expression

Government surveillance may violate the First Amendment when it interferes with protected activities like speech and assembly without a legitimate government interest, or when it targets specific groups for their political views to retaliate against them.¹⁵⁵ The uses described above—facial recognition at protests, AI monitoring of federal employees’ loyalty, and the collection and sharing of ALPR data—raise serious questions about the line between observing public conduct and chilling free expression. Courts have only begun to apply the Fourth and First Amendments to these technologies, and case law lags behind their deployment. Without clear legal limits, government surveillance has continued to expand.¹⁵⁶

¹⁵³ Ibid.

¹⁵⁴ *Zurcher v. Stanford Daily*, 436 U.S. 547 (1978).

¹⁵⁵ See Ashkhen Kazaryan, Is the Government Watching You? Surveillance and the First Amendment, Explained, *Freedom Forum*, April 13, 2026, <https://www.freedomforum.org/government-surveillance-first-amendment/>.

¹⁵⁶ Jay Stanley, “Machine Surveillance Is Being Super-Charged by Large AI Models,” American Civil Liberties Union, March 21, 2025, <https://www.aclu.org/news/privacy-technology/machine-surveillance-is-being-super-charged-by-large-ai-models>.

11. MEASURES EMPOWERING FREEDOM OF EXPRESSION

Given the challenges of AI-generated misinformation, institutions have placed a growing emphasis on media literacy—equipping the public to evaluate AI-generated content critically as an alternative to government restriction.¹⁵⁷ These educational efforts correlate with the constitutional preference for “counterspeech” over censorship.¹⁵⁸

President Trump’s April 2025 executive order, Advancing Artificial Intelligence Education for American Youth, directs federal agencies to promote and integrate AI literacy into K-12 curricula and educator training, and establishes a White House Task Force on AI Education and a Presidential AI Challenge. In June 2025, the White House announced an AI Education Pledge under which more than sixty organizations, including Microsoft, OpenAI, Google, Anthropic, and NVIDIA, committed to providing free tools, curricula, grants, and teacher training. In July 2025, the American Federation of Teachers launched the National Academy for AI Instruction with Microsoft, OpenAI, and Anthropic, aiming to train hundreds of thousands of K-12 educators.¹⁵⁹

The 2025 AI Action Plan calls for ensuring that AI protects free speech. Free expression advocates should nevertheless remain vigilant. The plan treats references to certain topics, such as diversity, equity, and inclusion (DEI) and climate change, as ideological bias, seeks to define neutrality, and emphasizes countering Chinese Communist Party (CCP) talking points. By presenting a single perspective as the standard of neutrality, the plan risks replacing one orthodoxy with another.¹⁶⁰

157 UNESCO, “AI Competency Framework for Students” (2024), <https://www.cedefop.europa.eu/en/tools/vet-toolkit-tackling-early-leaving/resources/unesco-ai-competency-framework-students>; National Association for Media Literacy Education (NAMLE), “AI Literacy Initiative,” September 9, 2024, <https://namle.org/ai-literacy-initiative/>.

158 American Federation of Teachers, “AFT to Launch National Academy for AI Instruction with Microsoft, OpenAI, Anthropic and United Federation of Teachers,” press release, July 8, 2025, <https://www.aft.org/press-release/aft-launch-national-academy-ai-instruction-microsoft-openai-anthropic-and-united>.

159 Exec. Order No. 14277, 90 Fed. Reg. 17519 (Apr. 28, 2025), <https://www.whitehouse.gov/presidential-actions/2025/04/advancing-artificial-intelligence-education-for-american-youth>; American Federation of Teachers, AFT to Launch National Academy for AI Instruction with Microsoft, OpenAI, Anthropic and United Federation of Teachers,” press release, July 8, 2025, <https://www.aft.org/press-release/aft-launch-national-academy-ai-instruction-microsoft-openai-anthropic-and-united>.

160 Exec. Order No. 14319, 90 Fed. Reg. 35389 (July 28, 2025), <https://www.whitehouse.gov/presidential-actions/2025/07/preventing-woke-ai-in-the-federal-government>.

12. MISCELLANEOUS

Two structural issues cut across the areas covered above: government control over access to advanced AI models, and US-China competition in AI.

12.1 Government Control of Advanced AI Models: Export Directives and the First Amendment

On June 12, 2026, the Commerce Department issued an export control directive barring all foreign nationals, inside or outside the United States, from accessing Anthropic’s two most advanced models, Fable 5 and Mythos 5. To comply, Anthropic disabled both models for all customers. Anthropic concluded that the alleged jailbreak vulnerability it understood to be the government’s concern was narrow and replicable with other publicly available models, and argued that the government should be able to block unsafe deployments only through a process that is “transparent, fair, clear, and grounded in technical facts.”¹⁶¹ The controls were lifted on June 30, 2026, following negotiations with the government, and Anthropic began restoring access on July 1.¹⁶²

The directive raises critical First Amendment questions. In *Reno v. ACLU* (1997), the Supreme Court struck down broad restrictions on online speech as incompatible with the First Amendment. In *TikTok v. Garland* (2025), by contrast, the court upheld a law requiring the divestiture or shutdown of an expressive platform on national security grounds. The Fable/Mythos directive could extend that precedent to AI models.¹⁶³

12.2 US-China AI Competition

Much of US federal AI policy during this period has been framed around competition with

¹⁶¹ Anthropic, “Statement on the US Government Directive to Suspend Access to Fable 5 and Mythos 5,” June 12, 2026, <https://www.anthropic.com/news/fable-mythos-access>.

¹⁶² Anthropic, “Redeploying Fable 5,” June 30, 2026, <https://www.anthropic.com/news/redeploying-fable-5>.

¹⁶³ *Reno v. ACLU*, 521 U.S. 844 (1997); *TikTok Inc. v. Garland*, 604 U.S. 56 (2025).

China, alongside the administration's focus on innovation. That framing has at times prompted scrutiny of the values embedded in AI models. The White House AI Action Plan directed the evaluation of Chinese models for alignment with CCP propaganda. In January 2026, the Bureau of Industry and Security formalized export controls on advanced AI chips, restricting their export to China and other countries and limiting the global diffusion of AI capabilities. Pursued without transparency or accountability, this geopolitical framing risks undue government intervention in AI development.¹⁶⁴

¹⁶⁴ OpenAI, "Announcing the Stargate Project," January 21, 2025, <https://openai.com/index/announcing-the-stargate-project/>; Bureau of Industry and Security, "Revision to License Review Policy for Advanced Computing Commodities," final rule, January 15, 2026, <https://www.federalregister.gov/documents/2026/01/15/2026-00789/revision-to-license-review-policy-for-advanced-computing-commodities>.

13. CONCLUSION

AI regulation in the United States is a patchwork of existing federal laws, executive actions, and state laws. Courts are working through a growing number of cases to determine how constitutional rights and limits apply to the development and use of AI.

Although the principles of the First Amendment are likely to extend to AI-generated content, how courts will apply them remains unsettled. Policymakers, legal scholars, and technology developers should work to ensure that AI's transformative capabilities expand free expression rather than constrain it. This approach recognizes that AI is both shaped by free speech protections and capable of strengthening the exercise of that right in the decades ahead.



**THE
FUTURE
OF
FREE
SPEECH**

www.futurefreespeech.org