



THE
FUTURE
OF
FREE
SPEECH

THAT VIOLATES MY POLICIES 2026

AI LAWS, CHATBOTS, AND
THE FUTURE OF EXPRESSION

DIRECTED BY:

Jordi Calvet-Bademunt,
Jacob Mchangama,
and Isabelle Anzabi

www.futurefreespeech.org

ACKNOWLEDGMENTS

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

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

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

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

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

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



PREFACE

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

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

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

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

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

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

THAT VIOLATES MY POLICIES 2026

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

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

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



ARTIFICIAL INTELLIGENCE & FREEDOM OF EXPRESSION IN THE EUROPEAN UNION

By

Jordi Calvet-Bademunt



JORDI CALVET-BADEMUNT

Jordi Calvet-Bademunt is a Senior Research Fellow at The Future of Free Speech at Vanderbilt University, where he leads the organization's research program on AI and freedom of expression. His work examines how AI governance and risk-management obligations affect fundamental rights, and his commentary has appeared in outlets including MSNBC, Lawfare, and Tech Policy Press. He is a member of the Freedom Online Coalition's Advisory Network, which advises the Coalition's more than forty member governments. Previously, Calvet-Bademunt spent nearly a decade working at the Organisation for Economic Co-operation and Development (OECD) and as an associate at leading European law firms. He holds advanced degrees from Harvard University and the College of Europe in Bruges, Belgium.

The author thanks Jacob Mchangama, Natalie Alkiviadou, and Alexander Hohlfeld for their valuable comments and suggestions. All remaining errors are my own.



ABSTRACT

This chapter explores how the laws and policies of the European Union (EU) affect free expression and access to information in the AI era. In 2024, the EU adopted the world's first cross-sector AI framework, the AI Act, and several of its provisions started to bite in 2026. We examine what the AI Act and the EU's rules on digital services, data protection, and copyright mean for free speech, across elections, hate speech, defamation, minors, explicit content, and surveillance. We analyze these issues in the context of the fundamental rules governing free expression, notably the EU Charter of Fundamental Rights and the European Convention on Human Rights.

The findings present a mixed picture. The EU has strong written protections for free expression, and most member states are robust democracies. The Charter and the Convention guarantee the right to receive and impart information, including ideas that offend, shock, or disturb. Some rules in the AI Act protect expression, notably the prohibitions on untargeted scraping of images to build facial recognition databases and on real-time biometric identification by law enforcement.

However, the EU and its member states have also adopted provisions that could restrict expression. In practice, Europe has been too quick to justify restrictions when speech is contentious, particularly on hate speech, and courts have often failed to check those excesses, especially online. In addition, provisions drafted in open-ended terms mix concrete risks with loosely defined concerns about disinformation and abstract harms to society. This vagueness invites misguided enforcement, legal uncertainty, and self-censorship.

The good news is that the Charter and the Convention, if applied with true commitment to our freedoms and pluralism, provide the guidance to ensure free expression and access to information are protected in this new era. Restrictions must be precise, pursue a legitimate aim, and be necessary and proportionate, adopted only after considering less restrictive alternatives.

1. INTRODUCTION

Generative artificial intelligence (AI) offers an unprecedented opportunity for freedom of expression and access to information. Never before has it been so easy to create content and learn about any topic, no matter how obscure. Language barriers have fallen dramatically, and creating art has become far more accessible. Dissidents and journalists across the world can use these tools to challenge established power. Hundreds of millions of us are embracing this technology.¹

Nevertheless, AI's promise comes with perils and inevitable trade-offs. The decisions of governments worldwide and of large tech companies will determine what fundamental rights look like in the AI era. This chapter focuses on how the European Union (EU) is approaching this new era.

The EU's decisions have an impact well beyond its borders, a phenomenon known as the "Brussels effect."² Companies often apply EU rules globally, and other countries replicate those rules in their legal frameworks. When Anthropic began watermarking Claude's output in August 2026 to comply with the AI Act, it applied the measure worldwide, saying it had no "durable way to scope it by region."³

The EU adopted the world's first cross-sector AI law, the AI Act, and many of its provisions are now applicable. Since August 2026, the European Commission has been able to enforce the rules for providers of general-purpose AI (GPAI) models, such as Anthropic, Google, and OpenAI, with fines of up to €15 million or 3 percent of worldwide annual turnover.⁴ That same month, the Commission designated ChatGPT as a very large online search engine under the Digital Services Act (DSA), Europe's online safety rulebook, bringing it within

1 Christine Zhenwei Qiang, Yan Liu, and He Wang, "Who on Earth Is Using Generative AI?," *Digital Transformation* (blog), World Bank, September 11, 2024, <https://blogs.worldbank.org/en/digital-development/who-on-earth-is-using-generative-ai->.

2 Anu Bradford, *The Brussels Effect: How the European Union Rules the World* (New York: Oxford University Press, 2020).

3 Anthropic, "How Claude's Text Watermarking Works," August 14, 2026, <https://www.anthropic.com/news/claude-text-watermark>. Anthropic and several other providers signed the EU Code of Practice on Transparency of AI-Generated Content in July 2026.

4 European Commission, "The Enforcement Framework of the AI Act," <https://digital-strategy.ec.europa.eu/en/policies/enforcement-ai-act>.

the scope of that law, as discussed in section 3.⁵ X had already encountered the DSA's impact on AI in January 2026, when the Commission opened formal proceedings over the deployment of Grok.⁶

The overall framework of freedom of expression in Europe will also affect the AI content we have access to. Europe is strongly committed to freedom of expression in the abstract but increasingly ready to restrict it in practice on contentious issues, particularly through broad definitions of hate speech and, more recently, disinformation. Major indices place EU countries in an enviable position on democracy and freedoms.⁷ Still, these freedoms require a permanent commitment, one that extends to the content we most despise.

This chapter covers EU law on freedom of expression, AI, and online services, together with the European Convention on Human Rights (the Convention), which, despite not being an EU document, shapes how fundamental rights are interpreted in the Union. Member state law appears only as illustration. Developments are covered through mid-September 2026.

5 European Commission, "Commission Designates ChatGPT, Reddit, Roblox under Digital Services Act," Press Release IP/26/1772, August 31, 2026, https://ec.europa.eu/commission/presscorner/detail/en/ip_26_1772.

6 European Commission, "Commission Investigates Grok and X's Recommender Systems under the Digital Services Act," Press Release IP/26/203, January 26, 2026, https://ec.europa.eu/commission/presscorner/detail/en/ip_26_203.

7 V-Dem Institute, *Democracy Report 2026: Unraveling the Democratic Era?*, 10th ed. (Gothenburg: University of Gothenburg, March 2026), https://www.v-dem.net/documents/75/V-Dem_Institute_Democracy_Report_2026_lowres.pdf; Freedom House, *Freedom in the World* 2026, <https://freedomhouse.org/report/freedom-world>.

2. GENERAL STANDARDS OF FREEDOM OF EXPRESSION

2.1 General Considerations

The EU is a political and economic union of twenty-seven countries—“member states,” in EU jargon—and the majority of them are ranked among the world’s stronger democracies. That record is not uniform, and there are warning signs. V-Dem’s 2026 report finds that “[a]utocratization in Europe affects seven EU member states.”⁸ Over the past few years, the most significant concern has been Hungary, which Freedom House ranks as “Partly Free”⁹ and which V-Dem classifies as an electoral autocracy.¹⁰ Over sixteen years, Viktor Orbán’s governments consolidated control of the media market, deployed spyware against journalists, subjected civil society to a “sovereignty protection” apparatus with investigative powers and no judicial oversight, and banned Pride marches.¹¹ Orbán lost power in the parliamentary election of April 2026,¹² opening the door to a potential change in Hungary’s situation.

In the EU, freedom of expression and access to information are enshrined in the Charter of Fundamental Rights (the Charter), which declares that “everyone has the right to freedom of expression,” including the “freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers.”¹³ The Charter binds the EU institutions, and it binds member states when they are implementing EU law.¹⁴

Freedom of expression is further protected by the Convention. The European Court of Human Rights (the Strasbourg Court), which applies the Convention, has clarified that

8 V-Dem Institute, *Democracy Report 2026*.

9 Freedom House, “After the Election: Revitalizing Hungarian Democracy, and Lessons from the Opposition Victory,” April 28, 2026, <https://freedomhouse.org/article/after-election-revitalizing-hungarian-democracy-and-lessons-opposition-victory>.

10 V-Dem Institute, *Democracy Report 2026*.

11 Jacob Mchangama, “Orbán’s War on Free Speech: The Receipts,” *The Bedrock Principle*, April 15, 2026, <https://www.bedrockprinciple.com/p/orbans-war-on-free-speech-the-receipts>.

12 Freedom House, “After the Election.”

13 Charter of Fundamental Rights of the European Union art. 11(1), 2012 O.J. (C 326) 391.

14 Charter of Fundamental Rights art. 51(1).

this protection extends even to information and ideas that “offend, shock or disturb.”¹⁵ The Convention is not an EU instrument, and the EU is not yet a party to it. Accession has been under negotiation for years, and in November 2025 the European Commission asked the Court of Justice of the European Union (the Luxembourg Court) to rule on whether the current draft accession agreement is compatible with the Treaties.¹⁶ The Convention is nonetheless influential in EU law. The Charter provides that, where it contains rights corresponding to Convention rights, their meaning and scope must be the same as under the Convention, although EU law may go further in protecting them.¹⁷ The Treaty on European Union recognizes Convention rights as general principles of EU law,¹⁸ and all twenty-seven member states have ratified the Convention.

Freedom of expression is not absolute. The Convention permits restrictions that are “prescribed by law” and “necessary in a democratic society”¹⁹ in pursuit of legitimate aims such as national security,²⁰ public safety,²¹ health or morals,²² or the rights and reputation of others.²³ The Charter is drafted in similar terms. Any limitation must be “provided for by law and respect the essence of those rights and freedoms,” must be proportionate, and must be “necessary and genuinely meet objectives of general interest recognised by the Union or the need to protect the rights and freedoms of others.”²⁴

15 *Handyside v. United Kingdom*, App. No. 5493/72, Eur. Ct. H.R. (December 7, 1976), para. 49, <https://hudoc.echr.coe.int/eng?i=001-57499>.

16 The Strasbourg Court should not be confused with the Luxembourg Court. The Luxembourg Court interprets EU law and settles disputes between EU institutions and member states; the Strasbourg Court was established by the Convention, outside the EU framework, and rules on compliance with it. On accession, see Request for an Opinion pursuant to Article 218(11) TFEU, Opinion 1/25, lodged by the European Commission on November 21, 2025, published in the Official Journal on January 26, 2026. The Luxembourg Court held the previous draft accession agreement incompatible with the Treaties in Opinion 2/13, ECLI:EU:C:2014:2454 (December 18, 2014).

17 Charter of Fundamental Rights art. 52(3).

18 Consolidated Version of the Treaty on European Union art. 6(3), 2012 O.J. (C 326) 13.

19 Convention for the Protection of Human Rights and Fundamental Freedoms art. 10(2), November 4, 1950, E.T.S. No. 5. Proportionality is a key element of such necessity: see, inter alia, *Lehideux and Isorni v. France*, App. No. 55/1997/839/1045, Eur. Ct. H.R. (September 23, 1998), <https://hudoc.echr.coe.int/eng?i=001-58245>.

20 See, inter alia, *Hadjianastassiou v. Greece*, App. No. 12945/87, Eur. Ct. H.R. (December 16, 1992), <https://hudoc.echr.coe.int/eng?i=001-57779>; *Stoll v. Switzerland*, App. No. 69698/01, Eur. Ct. H.R. (December 10, 2007), <https://hudoc.echr.coe.int/eng?i=001-83870>.

21 See, inter alia, *Leroy v. France*, App. No. 36109/03, Eur. Ct. H.R. (October 2, 2008), <https://hudoc.echr.coe.int/eng?i=002-1888>; *Stomakhin v. Russia*, App. No. 52273/07, Eur. Ct. H.R. (May 9, 2018, final October 8, 2018), <https://hudoc.echr.coe.int/eng?i=001-182731>; *Dmitriyevskiy v. Russia*, App. No. 42168/06, Eur. Ct. H.R. (October 3, 2017, final January 29, 2018), <https://hudoc.echr.coe.int/eng?i=001-177214>.

22 See, inter alia, *Bayev and Others v. Russia*, App. No. 67667/09, Eur. Ct. H.R. (June 20, 2017, final November 13, 2017), <https://hudoc.echr.coe.int/eng?i=001-174422>; *Müller and Others v. Switzerland*, App. No. 10737/84, Eur. Ct. H.R. (May 24, 1988), <https://hudoc.echr.coe.int/eng?i=001-57487>.

23 See, inter alia, *Féret v. Belgium*, App. No. 15615/07, Eur. Ct. H.R. (July 16, 2009), <https://hudoc.echr.coe.int/eng?i=002-1407>; *Seurot v. France* (dec.), App. No. 57383/00, Eur. Ct. H.R. (May 18, 2004), <https://hudoc.echr.coe.int/eng?i=002-4404>; *Vejdel and Others v. Sweden*, App. No. 1813/07, Eur. Ct. H.R. (February 9, 2012, final May 9, 2012), <https://hudoc.echr.coe.int/eng?i=001-109046>.

24 Charter of Fundamental Rights art. 52(1).

Outside Hungary, a key challenge to free expression comes from a growing body of criminal provisions that are applied expansively. In Germany, the Criminal Code includes an aggravated offense of insulting persons in political life, punished more severely than insults against ordinary citizens.²⁵ This is the opposite of what robust free speech standards require, since public officials and candidates should be expected to tolerate more criticism than ordinary citizens.²⁶ In France, the offense of “apology of terrorism” produced between 167 and 315 suspects per quarter from late 2023 to late 2024, according to figures the Ministry of Justice supplied to the National Assembly; more than a fifth of the first-instance decisions handed down in 2024 ended in acquittal.²⁷ In 2026, three UN special rapporteurs wrote to the French government warning that the offense is too vague to satisfy Article 19 of the International Covenant on Civil and Political Rights, and that it requires neither an intent to incite nor any likelihood that a terrorist act will follow.²⁸ Spain criminalizes the glorification of terrorism and the humiliation of its victims, which the Council of Europe’s Commissioner for Human Rights has faulted for its chilling effect and which has produced convictions of rappers, satirists, and ordinary social media users.²⁹ Spain also retains a crime of offending religious feelings, which private litigants have repeatedly invoked against satirists and media professionals.³⁰ To the current government’s credit, it has proposed repealing that offense along with the crime of insulting the Crown, though the reform has not yet been adopted.³¹

European courts have not consistently checked these trends, particularly on hate speech, as discussed in section 5. In *Sanchez v. France*, the Strasbourg Court upheld the criminal

25 Strafgesetzbuch [StGB] [Penal Code] § 188 (Ger.), https://www.gesetze-im-internet.de/englisch_stgb/englisch_stgb.html; see Alexander Hohlfeld, “Punishment Before Trial? ‘Hate Speech’ Raids in Germany,” *The Bedrock Principle*, October 24, 2025, <https://www.bedrockprinciple.com/p/punishment-before-trial-hate-speech>.

26 *Castells v. Spain*, App. No. 11798/85, judgment of April 23, 1992, para. 46, <https://hudoc.echr.coe.int/eng?i=001-57772>.

27 Assemblée nationale, Question écrite no. 9899 de Mme Gabrielle Cathala, “Nombre de procédures pour apologie de terrorisme et apologie de crime de guerre” [Number of proceedings for apology of terrorism and apology of war crimes], question published September 30, 2025, answer of the Ministry of Justice published May 12, 2026, <https://questions.assemblee-nationale.fr/q17/17-9899QE.htm> (Fr.). The offense is set out in Code pénal art. 421-2-5 (Fr.).

28 Communication OL FRA 16/2025 (January 12, 2026), from the Special Rapporteur on the promotion and protection of human rights and fundamental freedoms while countering terrorism, the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, and the Special Rapporteur on the situation of human rights defenders, <https://spcommreports.ohchr.org/TMResultsBase/DownloadPublicCommunicationFile?gld=30653>.

29 Código Penal art. 578 (Spain); Council of Europe Commissioner for Human Rights, “Misuse of Anti-Terror Legislation Threatens Freedom of Expression,” Human Rights Comment, December 4, 2018, <https://www.coe.int/en/web/commissioner/-/misuse-of-anti-terror-legislation-threatens-freedom-of-expression>.

30 Código Penal art. 525 (Spain).

31 “Spain to Scrap the Crime of Offending Religious Feelings Despite Opposition from Faith Groups,” *Evangelical Focus*, January 28, 2026, <https://evangelicalfocus.com/europe/33664/spain-to-scrap-the-crime-of-offending-religious-feelings-despite-opposition-from-faith-groups>.

conviction of a politician for failing to promptly delete unlawful third-party comments from his Facebook wall, extending intermediary-style liability to an individual social media user.³² More recently, in *Miladze v. Georgia*, the Court found no violation of Article 10 where a food-delivery courier was fined for insulting municipal officials in a TikTok video criticizing transport policy. Among the justifications the Court accepted was that the video had reached a wide audience, including minors, through algorithmic recommendation.³³ As one expert put it, “Strasbourg continues to proclaim robust protection for offensive political speech while simultaneously narrowing that protection in practice through elastic and ambiguous analyses.”³⁴

The Luxembourg Court has also endorsed limits that reduce the scope of information Europeans can access. In July 2026, it held that the EU’s prohibition on broadcasting Russia Today content applies to an individual who shared that content on a website financed entirely by voluntary donations. The Court held that the ban applies to anyone responsible, directly or indirectly, for making the prohibited content available, whether or not the activity is economic and whatever its extent or duration.³⁵ The Court did not appropriately address why such a broad reading was the least restrictive means of achieving the restriction’s purpose, and it gave journalists, researchers, and documentary makers who reproduce such material no clear guidance on their own legal exposure.³⁶

2.2 Considerations on AI

Two instruments commit the EU to protecting fundamental rights in the design and use of AI. The first is the AI Act, whose stated purpose is to improve the functioning of the internal market and promote human-centric and trustworthy AI “while ensuring a high level of protection of health, safety, fundamental rights enshrined in the Charter, including

32 *Sanchez v. France* [GC], App. No. 45581/15, Eur. Ct. H.R. (May 15, 2023), <https://hudoc.echr.coe.int/eng?i=001-224928>.

33 *Miladze v. Georgia*, App. No. 41585/23, Eur. Ct. H.R. (May 19, 2026), <https://hudoc.echr.coe.int/eng?i=001-250113>; European Court of Human Rights, Press Release ECHR 130 (2026), May 19, 2026.

34 Gionata Bouché, “The Politics of Provocation: The ECtHR’s Ruling in *Miladze v Georgia*,” *Verfassungsblog*, July 7, 2026, <https://verfassungsblog.de/miladze-echr/>. For a similar critique, see Babette De Naeyer, “Am I Allowed to Rant?: The Sour Aftertaste of *Miladze v. Georgia*,” *Strasbourg Observers*, June 30, 2026, <https://strasbourgobservers.com/2026/06/30/am-i-allowed-to-rantthe-sour-aftertaste-of-miladze-v-georgia/>.

35 Case C-67/25, *Traugott Ickerroth*, judgment of July 2, 2026; Court of Justice of the European Union, Press Release No. 94/26, July 2, 2026, <https://curia.europa.eu/site/upload/docs/application/pdf/2026-07/cp260094en.pdf>.

36 Natalie Alkiviadou, “When Exceptional Restrictions Stop Being Exceptional: The Court of Justice of the European Union and the Expanding Reach of the Russia Today Ban,” *The Bedrock Principle*, July 28, 2026, <https://www.bedrockprinciple.com/p/when-exceptional-restrictions-stop>.

democracy, the rule of law and environmental protection.”³⁷ Its recitals name freedom of expression and information among the Charter rights whose restriction is relevant in deciding whether an AI system is high-risk.³⁸ The second is the Council of Europe’s Framework Convention on Artificial Intelligence, the first binding international treaty on AI, which aims “to ensure that activities within the lifecycle of artificial intelligence systems are fully consistent with human rights, democracy and the rule of law.”³⁹ The EU ratified the Framework Convention in May 2026 and is so far its only party.⁴⁰ The treaty requires five ratifications, three of them by Council of Europe member states, before it enters into force.

At the Council of Europe level, the Steering Committee on Media and Information Society adopted a guidance note on the implications of generative AI for freedom of expression.⁴¹ The note proposes four criteria to assess whether AI-generated content should be protected: the degree of human agency behind the expression; the design choices made at each layer of the AI stack; the substance of what is conveyed; and the relationship between the human input and the output.⁴² This document is also notable for what it treats as risks. In addition to hallucinations, deepfakes, and voice cloning, it treats the “homogenisation of expression at scale, where unique or diverse voices risk being overshadowed” as a first-order concern.⁴³

Whether AI-generated content is protected speech will have a significant impact on our societies. By June 2026, ChatGPT reportedly had one billion monthly users.⁴⁴ According to a study by OpenAI researchers and a Harvard economist, about three-quarters of conversations in ChatGPT concerned practical guidance, information seeking, and writing.⁴⁵

37 Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) art. 1(1), 2024 O.J. (L, 2024/1689), as amended by Regulation (EU) 2026/1744, 2026 O.J. (L, 2026/1744).

38 Artificial Intelligence Act, recital 48.

39 Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law art. 1(1), opened for signature September 5, 2024, C.E.T.S. No. 225.

40 Council of Europe, “European Union Ratifies the Council of Europe Framework Convention on Artificial Intelligence,” May 15, 2026, <https://www.coe.int/en/web/artificial-intelligence/-/european-union-ratifies-the-council-of-europe-framework-convention-on-artificial-intelligence>.

41 Steering Committee on Media and Information Society (CDMSI), *Guidance Note on the Implications of Generative Artificial Intelligence for Freedom of Expression*, CDMSI(2025)15rev (adopted December 3–5, 2025; published March 2026), <https://rm.coe.int/cdmsi-2025-15rev-guidance-note-on-the-implications-of-generative-artif/488029df80>.

42 CDMSI, *Guidance Note*, para. 40.

43 CDMSI, *Guidance Note*, para. 52.

44 Matthew Chin, “ChatGPT Hits a Billion Monthly App Users Despite Souring Public AI Sentiment,” *CNBC*, June 12, 2026, <https://www.cnbc.com/2026/06/12/chatgpt-a-billion-monthly-app-users-despite-souring-public-ai-sentiment.html>.

45 OpenAI, “How People Are Using ChatGPT,” September 15, 2025, <https://openai.com/index/how-people-are-using-chatgpt/>; see also NBER Working Paper No. 34255.

Asking an AI model is increasingly a normal way of navigating questions ranging from trivial chores to complex research.

At the time of writing, neither the Strasbourg Court nor the Luxembourg Court has ruled on whether, or how, the right to freedom of expression protects AI-generated content. The most closely related case concerns copyright and is discussed in section 9.⁴⁶

Crucially, Article 11 of the Charter and Article 10 of the Convention protect the right to “receive” information and ideas, not merely to impart them.⁴⁷ Experts in the EU and the United States have argued that the listener’s standpoint offers a stronger basis for protecting generative AI content than the speaker’s standpoint.⁴⁸ Excluding AI outputs from protection precisely as people increasingly rely on them would shrink the pool of available information and perspectives and would give AI providers reasons to self-censor.

For these reasons, we believe that laws and policies that ban or otherwise limit access to AI content should be assessed under the ordinary standard for restrictions on expression. Any restrictions must be prescribed by law with enough precision for people to know what is expected of them, pursue a legitimate aim, and be necessary and proportionate to that aim. Vague and broad restrictions on AI-generated content would place satire, artistic creation, and public discourse at risk.

46 Case C-250/25, *Like Company v. Google Ireland Ltd.*, 2025 O.J. (C, C/2025/3039); hearing before the Grand Chamber, March 10, 2026.

47 Charter of Fundamental Rights art. 11(1); Convention art. 10(1).

48 Marco Bassini, “Speech Without a Speaker: Constitutional Coverage for Generative AI Output?,” *European Constitutional Law Review* 21, no. 3 (2025): 375–411, <https://doi.org/10.1017/S1574019625100771>; see also Eugene Volokh, Mark A. Lemley, and Peter Henderson, “Freedom of Speech and AI Output,” *Journal of Free Speech Law* 3, no. 2 (2023): 651–59, <https://www.journaloffreespeechlaw.org/volokhlemleyhenderson.pdf>.

3. AI LEGISLATION LANDSCAPE

In the EU, the main framework governing AI is the AI Act. It is a regulation aimed at improving the functioning of the internal market by preventing a patchwork of national laws while pursuing the fundamental-rights aims mentioned in section 2.⁴⁹ The AI Act sits alongside the DSA, the General Data Protection Regulation (GDPR), EU copyright law, and other digital laws; it is complemented by codes of practice and Commission guidelines; and it is supplemented by some national rules.

3.1 The AI Act

The AI Act is built around gradations of risk. Some uses are treated as unacceptable and banned outright. Others fall into a “high-risk” category and face obligations on risk management, data governance, documentation, and human oversight. Systems posing only limited risk are subject mainly to transparency duties, including letting users know when they are dealing with AI rather than a human, and marking or labeling synthetic media.⁵⁰

Generative AI was not specifically covered by the Commission’s original AI Act proposal. Provisions on GPAI models—models trained on very large amounts of data that can perform a wide range of tasks—were introduced later in the process, after the release of ChatGPT in November 2022.⁵¹ Providers of GPAI models must keep technical documentation, put in place a policy to comply with EU copyright law, and publish a sufficiently detailed summary of the content used to train the model. Models deemed to carry “systemic risk”—a category that captures, among others, models trained using large amounts of computing power—face a further set of duties, including model evaluations, assessment and mitigation of systemic risk, reporting of serious incidents, and cybersecurity protection.⁵²

49 Artificial Intelligence Act art. 1(1).

50 Artificial Intelligence Act arts. 5, 6 and 50.

51 Oskar J. Gstrein, Noman Haleem, and Andrej Zwitter, “General-Purpose AI Regulation and the European Union AI Act,” *Internet Policy Review* 13, no. 3 (2024), <https://doi.org/10.14763/2024.3.1790>. The Commission’s proposal of 21 April 2021 contained no provisions on general-purpose AI models; the European Parliament’s negotiating position of 14 June 2023 was the first version of the text to address them.

52 Artificial Intelligence Act arts. 51–55.

Some of the concerns behind systemic risk provisions are well founded. Over the summer of 2026, AI models left their evaluation environments and intruded into the production systems of real companies.⁵³ Cybersecurity risks of that kind are concrete and demonstrable, and we need to tackle them. However, the current provisions mix real issues, like cybersecurity,⁵⁴ with vaguely defined concerns like disinformation and speculative harms to society. This mix can result in misguided enforcement, legal uncertainty, and self-censorship.

A significant part of the operative detail of the AI Act is developed in accompanying documents. The GPAI Code of Practice, published in July 2025, was drawn up by independent experts and gives providers a way of demonstrating compliance. Adherence is voluntary, but the Commission and the AI Board have confirmed the Code as an adequate tool for that purpose, and more than twenty providers—among them Amazon, Anthropic, Google, Microsoft, Mistral AI, and OpenAI—have signed.⁵⁵ Meta declined,⁵⁶ and xAI signed only the safety and security chapter.⁵⁷ A separate Code of Practice on Transparency of AI-Generated Content, published in June 2026, performs the same function for the marking and labeling duties in Article 50.⁵⁸ The Commission subsequently approved guidelines on those duties.⁵⁹ None of these instruments binds a court, but national market surveillance authorities and the AI Office can be expected to follow them.

Parts of the AI Act are already applicable. The Commission can enforce the GPAI rules directly, with fines of up to €15 million or 3 percent of worldwide annual turnover, whichever

53 See Hugging Face, “Anatomy of a Frontier Lab Agent Intrusion: A Technical Timeline of the July 2026 Incident,” July 27, 2026, <https://huggingface.co/blog/agent-intrusion-technical-timeline>. Other cases were disclosed over the same weeks: see Anthropic, “Investigating Three Real-World Incidents in Our Cybersecurity Evaluations,” July 30, 2026, <https://www.anthropic.com/news/investigating-incidents-cybersecurity-evals>.

54 Artificial Intelligence Act, recital 110, which lists among systemic risks “offensive cyber capabilities, such as the ways in vulnerability discovery, exploitation, or operational use can be enabled.” The Safety and Security chapter of the General-Purpose AI Code of Practice sets out practices for assessing and mitigating such risks. See European Commission, “The General-Purpose AI Code of Practice,” <https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai>.

55 European Commission, “The General-Purpose AI Code of Practice,” published July 10, 2025, <https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai> (signatory list as consulted in September 2026).

56 Will Shanklin, “Meta Says It Won’t Sign the EU’s AI Code of Practice,” *Engadget*, July 18, 2025, <https://www.engadget.com/ai/meta-says-it-wont-sign-the-eus-ai-code-of-practice-190132690.html>.

57 European Commission, “The General-Purpose AI Code of Practice.”

58 European Commission, “Code of Practice on Transparency of AI-Generated Content,” June 10, 2026, <https://digital-strategy.ec.europa.eu/en/policies/code-practice-ai-generated-content>.

59 European Commission, “Guidelines on Transparency Obligations for Providers and Deployers of AI Systems,” approved July 20, 2026, <https://digital-strategy.ec.europa.eu/en/library/guidelines-transparency-obligations-providers-and-deployers-ai-systems>; European Commission, Press Release IP/26/1653 (July 20, 2026).

is higher.⁶⁰ The Commission has sent the first requests for information to companies on “cybersecurity, safety and copyright compliance.”⁶¹ After a reform adopted in 2026 under the name of “AI Omnibus,” the Act also gives the AI Office exclusive competence over AI systems built on a general-purpose model where the system and the model are offered by the same provider, and over AI systems integrated into very large online platforms and search engines regulated under the DSA.⁶²

The AI Omnibus also added a new prohibition, applicable from December 2, 2026, on AI systems that generate or manipulate child sexual abuse material (CSAM) or nonconsensual intimate imagery (NCII). The prohibition applies where such output is a reasonably foreseeable and reproducible outcome achievable without significant technical modification, and where the system lacks adequate technical safeguards.⁶³ The AI Omnibus also postponed the obligations for high-risk systems until December 2027 or August 2028, depending on the system. Article 50 transparency duties benefit from a grace period, until December 2, 2026, for marking systems already on the market.⁶⁴ Finally, the AI literacy obligation in Article 4 was softened. Providers and deployers no longer have to ensure an adequate level of AI literacy among their staff, but only to take measures supporting its development.⁶⁵

This chapter assesses the AI Act’s impact on freedom of expression. Some provisions protect expression, notably the limits on biometric surveillance that safeguard public engagement (section 10). Others put it at risk, particularly the obligations on providers of GPAI models with systemic risk. The ultimate result depends, in part, on how the law is interpreted and applied. As of September 2026, no court has ruled on the Act, and no enforcement decision has been adopted. Still, the GPAI Code of Practice suggests that a restrictive direction cannot be excluded.

60 European Commission, “Enforcement Framework of the AI Act.”

61 Maximilian Henning, “EXCLUSIVE: EU Orders Leading AI Labs to Detail Security Practices,” *Euractiv*, August 26, 2026, <https://www.euractiv.com/news/exclusive-eu-orders-leading-ai-labs-to-detail-security-practices/>.

62 Regulation (EU) 2026/1744 of the European Parliament and of the Council amending Regulation (EU) 2024/1689, 2026 O.J. (L, 2026/1744), <http://data.europa.eu/eli/reg/2026/1744/oj>; see White & Case, “EU AI Omnibus Enters into Force, Amending the AI Act,” <https://www.whitecase.com/insight-alert/eu-ai-omnibus-enters-force-amending-ai-act>.

63 Regulation (EU) 2026/1744, amending Artificial Intelligence Act art. 5; Patrick Van Eecke and Bartholomäus Regenhart, “Digital AI Omnibus Delays Key Deadlines, Introduces New Rules,” *Cooley*, August 3, 2026, <https://cdp.cooley.com/digital-ai-omnibus-delays-key-deadlines-introduces-new-rules/>. Deployers are caught only where they intentionally use a system for that purpose.

64 Orrick, “EU AI Act Update: Digital Omnibus Finalizes 8 Compliance Changes,” July 2026, <https://www.orrick.com/en/Insights/2026/07/EU-AI-Act-Update-Digital-Omnibus-Finalizes-8-Compliance-Changes>.

65 Regulation (EU) 2026/1744, amending Artificial Intelligence Act art. 4.

The AI Act was not, in the end, complemented by the proposed AI Liability Directive, which the Commission withdrew in 2025.⁶⁶ The EU did, however, update its product liability regime. A 2024 directive treats software—including standalone software, and therefore AI systems—as a “product,” so that manufacturers and providers face liability without fault for defects, including defects that emerge through updates or continued learning after a product reaches the market. It also eases the claimant’s burden of proof.⁶⁷

The Framework Convention on AI (section 2) binds parties in respect of activities carried out by public authorities and by private actors acting on their behalf. As regards the private sector more broadly, parties must address risks and impacts but retain considerable latitude in how they do so.⁶⁸

3.2 The DSA

Generative AI also falls within the scope of the DSA. The DSA requires very large online platforms (VLOPs) and very large online search engines (VLOSEs) to assess and mitigate “systemic risks” arising from their services. These risks are not identical to the systemic risks in the AI Act, but they overlap substantially and are also worryingly vague. The overlap between the two laws is not theoretical. In August 2026, the Commission designated ChatGPT as a VLOSE, bringing an AI product within the remit of the DSA.⁶⁹ The Commission had already opened proceedings against X in January 2026 over risks arising from the deployment of Grok on that platform.⁷⁰

The DSA grants hosting services a liability exemption for third-party content, provided they act expeditiously to remove illegal material once notified. This exemption is highly unlikely to cover content autonomously generated by an AI model. In May 2026, a Munich court held

66 Caitlin Andrews, “European Commission Withdraws AI Liability Directive from Consideration,” *IAPP*, February 12, 2025, <https://iapp.org/news/a/european-commission-withdraws-ai-liability-directive-from-consideration>.

67 Directive (EU) 2024/2853 of the European Parliament and of the Council of 23 October 2024 on liability for defective products and repealing Council Directive 85/374/EEC, 2024 O.J. (L, 2024/2853); Christina Kiefer and Laure Herlitz, “Liability for Software under the New European Product Liability Directive,” *International Bar Association*, April 30, 2025, <https://www.ibanet.org/European-Product-Liability-Directive-liability-for-software>.

68 Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, C.E.T.S. No. 225, arts. 1, 3; see section 2 above.

69 European Commission, “Commission Designates ChatGPT.” Reddit and Roblox were designated very large online platforms on the same date.

70 European Commission, “Commission Investigates Grok and X’s Recommender Systems.”

Google directly liable for false statements generated by its AI Overviews feature (section 6).⁷¹ If liability extends to AI output generated from third-party content, even when the origin of such content is disclosed, users can expect access to a much narrower range of opinions and perspectives.

The Luxembourg Court has narrowed the hosting exemption in two recent judgments. In *Russmedia*, it held that the exemption does not relieve a marketplace operator of its obligations as a data controller. Worryingly, it found that appropriate measures may require the operator to identify advertisements containing sensitive data before publication and to verify that the person posting them is the person they concern.⁷² In another case, the Court held that an operator whose algorithm determines, in its own interest and beyond mere categorization and indexing, “under what conditions, how and in which order of priority” information is broadcast exercises control over it.⁷³ As a result, it falls outside the exemption and loses the protection against general monitoring obligations.⁷⁴

3.3 National Legislation

Some EU member states have also begun to legislate on AI. In 2025, Italy adopted a law that created a new criminal offense of unlawfully disseminating AI-generated or altered content if it causes unjust harm, punishable by one to five years’ imprisonment, with no express exception for satire, parody, or art.⁷⁵ Spain’s draft organic law on the good use and governance of AI would make failure to label AI-generated content a serious infringement carrying significant fines.⁷⁶ Denmark has proposed amendments to its Copyright Act to give every person a right over AI-generated reproductions of their voice and appearance, though

71 Landgericht München I, 26 O 869/26, judgment of May 28, 2026; see Giovanni Vetugno, “Who Speaks When the Algorithm Speaks? A German Ruling on AI Overviews,” *Oxford Business Law Blog*, July 16, 2026, <https://blogs.law.ox.ac.uk/oblb/blog-post/2026/07/who-speaks-when-algorithm-speaks-german-ruling-ai-overviews>.

72 Case C-492/23, *Russmedia Digital and Inform Media Press*, judgment of December 2, 2025 (Grand Chamber).

73 Joined Cases C-188/24 and C-190/24, *WebGroup Czech Republic and NKL Associates and Coyote System*, judgment of June 16, 2026 (Grand Chamber), paras. 110–112 and 122; Court of Justice of the European Union, Press Release No. 87/26, June 16, 2026, <https://curia.europa.eu/site/upload/docs/application/pdf/2026-06/cp260087en.pdf>.

74 *WebGroup Czech Republic*, paras. 110–112 and 122.

75 Legge 23 settembre 2025, n. 132, “Disposizioni e deleghe al Governo in materia di intelligenza artificiale” [Provisions and delegations to the Government on artificial intelligence], art. 4(4), *Gazzetta Ufficiale della Repubblica Italiana*, no. 223 (September 25, 2025), in force October 10, 2025, [https://www.gazzettaufficiale.it/eli/id/2025/09/25/25G00143/sg; Codice penale art. 612-quater \(It.\)](https://www.gazzettaufficiale.it/eli/id/2025/09/25/25G00143/sg; Codice penale art. 612-quater (It.)).

76 Proyecto de Ley Orgánica para el buen uso y la gobernanza de la Inteligencia Artificial [Draft Organic Law on the Good Use and Governance of Artificial Intelligence], approved by the Council of Ministers and transmitted to the Congress of Deputies in June 2026 (Spain); see finreg360, “El Congreso inicia el trámite del proyecto de ley orgánica para el buen uso y la gobernanza de la inteligencia artificial” [Congress begins consideration of the draft organic law on the good use and governance of artificial intelligence], <https://finreg360.com/alerta/el-congreso-inicia-el-tramite-del-proyecto-de-ley-organica-para-el-buen-uso-y-la-gobernanza-de-la-inteligencia-artificial/>.

caricature, satire, parody, and pastiche would in principle be exempt.⁷⁷ These national laws and initiatives give some indication of how free expression may be restricted in the AI era.

⁷⁷ Sofia Karttunen, “The Danish Approach to Copyright and Deepfakes: A Model for the EU?,” EPRS At a Glance PE 782.611 (January 2026), [https://www.europarl.europa.eu/RegData/etudes/ATAG/2026/782611/EPRS_ATA\(2026\)782611_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/ATAG/2026/782611/EPRS_ATA(2026)782611_EN.pdf); Alma Eggers, “Copyrighting Voice and Image: How Denmark’s New Deepfake Law Challenges Traditional Personality Rights in Europe,” *Verfassungsblog*, March 26, 2026, <https://verfassungsblog.de/copyrighting-voice-and-image/>.

4. ELECTION AND POLITICAL CONTENT

4.1 Background

AI is changing political and election communication in significant ways, some concerning, some positive. On the one hand, it can be used to fabricate news, impersonate candidates, and manufacture the appearance of grassroots support. On the other, it lowers the cost of producing political communication, which in principle favors candidates, campaigners, and civil society organizations that do not have large budgets.

At the same time, AI does not change the fact that democratic societies should start from the premise that political speech receives the highest protection. Citizens must be free to criticize their governments and to argue for change. The Strasbourg Court has consistently held that there is little scope under Article 10 of the Convention for restrictions on political speech or on debate about questions of public interest.⁷⁸ The first response to the risks should therefore rely on nonrestrictive measures, including greater transparency, media literacy, and a diverse media landscape, as the high-level expert group convened by the Commission recommended in 2018.⁷⁹

The EU has adopted measures of that kind. As section 10 explains, the AI Act restricts several uses of AI for mass surveillance, which protects political expression, including participation in demonstrations. The DSA requires providers to give users a statement of reasons whenever they remove content or restrict its visibility, which makes moderation decisions visible and contestable.⁸⁰

⁷⁸ European Court of Human Rights, *Guide on Article 10 of the European Convention on Human Rights: Freedom of Expression*, para. 146, <https://rm.coe.int/guide-on-article-10-freedom-of-expression-eng/native/1680ad61d6>.

⁷⁹ European Commission, *Final Report of the High Level Expert Group on Fake News and Online Disinformation*, March 12, 2018, <https://digital-strategy.ec.europa.eu/en/library/final-report-high-level-expert-group-fake-news-and-online-disinformation>.

⁸⁰ Regulation (EU) 2022/2065 on a Single Market for Digital Services (Digital Services Act) art. 17, 2022 O.J. (L 277) 1.

4.2 Deepfakes

The EU also aims to respond to deceptive content by mandating transparency. Under Article 50 of the AI Act, providers must mark synthetic content in a machine-readable way, and deployers of AI systems that generate or manipulate deepfakes must disclose that the content is artificial.⁸¹

The Act defines a deepfake as content resembling “existing persons, objects, places, entities or events,” and the Commission guidelines interpret “existing” as requiring realism, so that unrealistic or wholly invented depictions are not covered by the obligation.⁸² AI-generated text published to inform the public on matters of public interest is exempt where it has undergone genuine editorial control, understood as a deliberate, substantive examination by a competent person, with an identified natural or legal person bearing responsibility for the publication.⁸³ Importantly, the Act provides that a breach of Article 50 does not itself affect the legality of the content. This means that an unlabeled deepfake does not automatically become “illegal content” that platforms must remove under the DSA.⁸⁴

The protection of satire is weaker. A deepfake that forms part of an artistic, creative, satirical, or fictional work is not exempt. The obligation is only attenuated, requiring a disclosure in a manner that does not hamper the display or enjoyment of the work.⁸⁵ In addition, the content must “evidently” fall within one of the listed categories (artistic, creative, satirical, or fictional work).⁸⁶ In practice, this puts the burden on the speaker to make the joke obvious enough to satisfy a regulator, reducing the ambiguity on which satire, parody, and humor frequently depend. The stakes are high. A breach of Article 50 carries fines of up to €15 million or 3 percent of worldwide annual turnover, and the obligations bind deployers, including individuals whenever they are acting outside a purely personal, nonprofessional activity.⁸⁷

81 Artificial Intelligence Act art. 50(2) and (4).

82 Artificial Intelligence Act art. 3(60); European Commission, Guidelines on Transparency, para. 115.

83 European Commission, Guidelines on Transparency, paras. 134–38.

84 Artificial Intelligence Act recital 136; Nicolaj Feltes, “Article 50 AI Act: Do the Transparency Provisions Improve Upon the Commission’s Draft?,” *Journal of Intellectual Property, Information Technology and Electronic Commerce Law* 16 (2025): paras. 86, 90, <https://www.jipitec.eu/jipitec/article/view/438>.

85 Artificial Intelligence Act art. 50(4); European Commission, Guidelines on Transparency, paras. 122–24.

86 Feltes, “Article 50 AI Act,” paras. 62–65.

87 Artificial Intelligence Act arts. 3(4) and 99(4); on the applicable dates, see section 3 above.

These measures are accompanied by more restrictive ones at the national level. In 2018, France adopted a law empowering judges to order the removal of “fake news” within forty-eight hours during election campaigns.⁸⁸ Croatia, Czechia, and Hungary, among others, have criminal provisions covering the deliberate dissemination of disinformation that threatens peace or public order.⁸⁹ The former UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression warned that disinformation is an “extraordinarily elusive concept to define in law,” apt to give authorities “excessive discretion to determine what is disinformation, what is a mistake, what is truth.”⁹⁰

4.3 Risk-Based Provisions

In the AI Act, systemic-risk obligations raise the most serious concern. Providers of GPAI models with systemic risk must assess and mitigate that risk, and the Act defines it as a risk specific to a model’s high-impact capabilities with “actual or reasonably foreseeable negative effects on public health, safety, public security, fundamental rights, or the society as a whole, that can be propagated at scale across the value chain.”⁹¹ The expression “society as a whole” gives wide discretion to regulators and companies. This raises the risk of misguided enforcement, and of companies restricting lawful speech to avoid liability.

The GPAI Code of Practice does little to ease this concern. Although not mandatory, as explained in section 3, it is influential and carries the backing of most major providers, and it identifies “false” and “radicalizing” content among the risks that companies should consider mitigating.⁹²

The DSA uses similar language to the AI Act, and it also applies to generative AI services following ChatGPT’s designation as a VLOSE (section 3). It has already shown how such

88 Loi n° 2018-1202 du 22 décembre 2018 relative à la lutte contre la manipulation de l’information [Law 2018-1202 of December 22, 2018, Relating to the Fight Against the Manipulation of Information], Journal Officiel de la République Française, December 23, 2018, <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000037847559> (Fr.).

89 Konrad Bleyer-Simon, Elda Brogi, Iva Nenadić, and Teona Nesović, “Policies to Tackle Disinformation in EU Member States – Part II,” European Digital Media Observatory and Centre for Media Pluralism and Media Freedom (2022): 11, <https://edmo.eu/wp-content/uploads/2022/07/Policies-to-tackle-disinformation-in-EU-member-states-%E2%80%93-Part-II.pdf>.

90 Human Rights Council, Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, “Disease Pandemics and the Freedom of Opinion and Expression,” A/HRC/44/49 (April 23, 2020), <https://www.ohchr.org/en/documents/thematic-reports/ahrc4449-disease-pandemics-and-freedom-opinion-and-expression-report>, quoted in Natalie Alkiviadou, “Prison for Fake News,” *Verfassungsblog*, July 12, 2024, <https://verfassungsblog.de/prison-for-fake-news>.

91 Artificial Intelligence Act art. 3(65); see also arts. 51 and 55.

92 European Commission, “Safety and Security,” in *General-Purpose AI Code of Practice*, 34, <https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai>.

provisions can be misused. Former Commissioner Thierry Breton sent companies letters that conflated illegal content with disinformation,⁹³ and warned X about its systemic-risk obligations while cautioning Elon Musk against “harmful content” ahead of an interview between Musk and then-candidate Donald Trump on the platform.⁹⁴

The Commission’s guidelines on mitigating systemic risks to electoral processes go further. They recommend that companies adapt their terms and conditions “to significantly decrease the reach and impact of generative AI content that depicts disinformation or misinformation.”⁹⁵ Given how contested the boundaries of disinformation are, and given that false speech is in principle protected,⁹⁶ a recommendation that platforms suppress the reach of lawful content is difficult to reconcile with Article 11 of the Charter.

It is doubtful whether systemic-risk provisions in their current form satisfy Article 52 of the Charter. Researchers have argued that these obligations “pose unique problems when seen from the popular three-pronged test used by apex courts around the world to assess restrictions on freedom of expression,”⁹⁷ and have identified a “fundamental tension” between a risk-based approach and the principle of legality.⁹⁸ On proportionality, they note that the assessment is outsourced to companies, and that it is carried out against the risks being mitigated rather than against the legitimate aims the legislation pursues.⁹⁹

Other experts argue that “any attempts [through the DSA systemic risk provisions] to prohibit or suppress specific expression that is legal would arguably violate the rule of law because of the insufficient legal basis.”¹⁰⁰ The AI Act’s systemic-risk obligations are built on the same model as the DSA, and the same logic applies to them. The Commission and

93 Access Now, “When Tackling Illegal Online Content Related to Conflict in Gaza, the Rule of Law Matters,” press release, October 18, 2023, <https://www.accessnow.org/press-release/dsa-gaza-online-content>.

94 Thierry Breton (@ThierryBreton), “With great audience comes greater responsibility #DSA,” X, August 12, 2024, <https://x.com/Thierry-Breton/status/1823033048109367549>.

95 Guidelines for Providers of Very Large Online Platforms and Very Large Online Search Engines on the Mitigation of Systemic Risks for Electoral Processes Pursuant to Article 35(3) of Regulation (EU) 2022/2065, 2024 O.J. (C 3014), para. 40 (April 26, 2024), <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52024XC03014>.

96 *Salov v. Ukraine*, App. No. 65518/01, Eur. Ct. H.R. (September 6, 2005, final December 6, 2005), para. 113, <https://hudoc.echr.coe.int/eng?i=001-70096>.

97 Agustina Del Campo, Nicolás Zara, and Ramiro Álvarez Ugarte, “Are Risks the New Rights? The Perils of Risk-Based Approaches to Speech Regulation,” Center for Studies on Freedom of Expression, *Artículo de investigación* 64 (February 25, 2025): 1, https://www.palermo.edu/Archivos_content/2025/cele/febrero/2025_02_01_Are_Risk_CELE.pdf.

98 Del Campo, Zara, and Álvarez Ugarte, “Are Risks the New Rights?,” 18.

99 Del Campo, Zara, and Álvarez Ugarte, “Are Risks the New Rights?,” 19.

100 Martin Husovec, *Principles of the Digital Services Act* (Oxford: Oxford University Press, 2024), 335.

the Luxembourg Court should state clearly that the DSA and the AI Act do not empower regulators to restrict legal content. Until they do, the risk of misguided enforcement—and of companies restricting lawful speech in order to stay on the safe side—remains.

The AI Act also classifies as high-risk any AI system “intended to be used for influencing the outcome of an election or referendum or the voting behaviour of natural persons,” while excluding systems to whose output voters are not directly exposed, such as campaign logistics tools.¹⁰¹ Those obligations were postponed by the 2026 reform and now apply from December 2, 2027.¹⁰²

Although the Act requires intent, a Dutch study shows that this provision presents risks for free expression and access to information. In October 2025, ahead of the Dutch parliamentary elections, the data protection authority tested four widely used chatbots—ChatGPT, Gemini, Mistral, and Grok—by building 3,000 fictional voter profiles and asking each system for a top-three recommendation. The results were heavily skewed. The authority described a “vacuum-cleaner effect” in which profiles were pulled toward the edges—the GroenLinks-PvdA on the left and the PVV on the right—so that “the political middle ground virtually disappears.”¹⁰³ The pattern held across every system tested. As a result, the authority called on providers “to prevent their systems from being used for voting advice,” and took the position that “AI systems that provide voting advice should meet the strict requirements that apply to high-risk systems under the AI Act,” citing the provision on influencing an election mentioned above.¹⁰⁴ Asking providers to stop their systems from answering questions about how to vote requires companies to withhold lawful political information from citizens who have sought it out, directly impacting their right to access information.

4.4 Other Initiatives

The EU is still adding to this landscape. In November 2025, the Commission and the

¹⁰¹ Artificial Intelligence Act Annex III, point 8(b).

¹⁰² Regulation (EU) 2026/1744; see section 3 above.

¹⁰³ Autoriteit Persoonsgegevens, Department for the Coordination of Algorithmic Oversight, *RAN Special: AI Chatbots as Voting Aid* (October 2025), 3, <https://www.autoriteitpersoonsgegevens.nl/en/current/ap-warns-chatbots-give-biased-voting-advice>.

¹⁰⁴ Autoriteit Persoonsgegevens, *AI Chatbots as Voting Aid*, 4.

High Representative issued a joint communication on a “European Democracy Shield.”¹⁰⁵ It commits the Commission to preparing “a dedicated guidance on the fair, transparent, human-centred and responsible use of AI in electoral processes” and to updating the DSA elections toolkit, and it includes measures such as reinforced funding for independent and local journalism.¹⁰⁶ The European Parliament’s special committee adopted its own report in June 2026, and a plenary vote was expected in September 2026.¹⁰⁷ Advocates should pay close attention to this development to see whether the Shield strengthens democratic resilience or promotes restrictions on information and expression.

¹⁰⁵ European Commission and High Representative of the Union for Foreign Affairs and Security Policy, Joint Communication, “European Democracy Shield: Empowering Strong and Resilient Democracies,” November 12, 2025, https://commission.europa.eu/document/download/2539eb53-9485-4199-bfdc-97166893ff45_en.

¹⁰⁶ European Commission and High Representative, “European Democracy Shield,” 12.

¹⁰⁷ European Parliament, Special Committee on the European Democracy Shield, report adopted June 23, 2026; European Parliament, Legislative Train Schedule, “European Democracy Shield,” last updated June 20, 2026, <https://www.europarl.europa.eu/legislative-train/package-european-democracy-action-plan/file-european-democracy-shield>.

5. HATE SPEECH

The EU and its member states combat hate speech through a range of instruments, and those instruments have implications for AI-generated content. Hate speech is abhorrent and can cause pain. It is also a slippery concept that reaches into contested sociopolitical argument, and criminal sanctions risk turning those convicted into “free speech martyrs.”¹⁰⁸ Where possible, counterspeech is the better answer for freedom of expression.¹⁰⁹

Europe has seen firsthand how hate-speech rules can misfire. No recent controversy shows the difficulty more clearly than the conflict between Israel and Palestine, where the same words can be heard as a demand for Palestinian self-determination or a call for the destruction of Israel, depending on who is listening. Germany is the sharpest illustration. An Israeli-German citizen was arrested five times for demonstrating with a sign reading “As a Jew and Israeli, stop the genocide in Gaza”; the incidents were reportedly logged as hate speech and counted in national antisemitism statistics.¹¹⁰ In 2024, the domestic intelligence service designated Jüdische Stimme, a Jewish organization critical of Israeli policy, as extremist.¹¹¹

AI-generated hate speech has already resulted in calls for investigations, and investigations themselves. In July 2025, a Polish minister announced that the government would report X to the European Commission after Grok produced antisemitic output and offensive comments about Polish politicians, including Prime Minister Donald Tusk.¹¹² In November 2025, Grok posted a Holocaust denial message. French ministers referred the output to

108 Aleksandra Gliszczynska-Grabias, “Penalising Anti-Semitism in Poland: Creating ‘Free Speech Martyrs’?,” *OxHRH Blog*, July 18, 2017, <https://ohrh.law.ox.ac.uk/penalising-anti-semitism-in-poland-creating-free-speech-martyrs/>.

109 The Future of Free Speech and the Dangerous Speech Project, *A Toolkit on Using Counterspeech to Tackle Online Hate Speech*, <https://futurefreespeech.org/a-toolkit-on-using-counterspeech-to-tackle-online-hate-speech/>.

110 Shira Li Bartov, “How Germany’s Unusual Approach to Fighting Antisemitism Is Ensnaring Jews Who Are Critical of Israel,” *Jewish Telegraphic Agency*, March 4, 2026, <https://www.jta.org/2026/03/04/global/how-germanys-unusual-approach-to-fighting-antisemitism-is-ensnaring-jews-who-are-critical-of-israel>.

111 Bartov, “How Germany’s Unusual Approach.”

112 “Poland to Report Musk’s Chatbot Grok to EU for Offensive Comments, Including on PM Tusk,” *Malay Mail*, July 9, 2025, <https://www.malaymail.com/news/tech-gadgets/2025/07/09/poland-to-report-musks-chatbot-grok-to-eu-for-offensive-comments-including-on-pm-tusk/183395>.

prosecutors, and the Paris public prosecutor's office added it to an existing investigation of X.¹¹³ That office announced that it was seeking charges against X and Elon Musk, including for denial of crimes against humanity, complicity in the distribution of CSAM, and the dissemination of nonconsensual intimate images.¹¹⁴ This appears to be the first criminal proceeding in the EU arising from hate speech generated by an AI system.

Restrictions on hate speech must comply with the right to freedom of expression. However, the case law of the Strasbourg Court has failed to establish a stable doctrine on this issue. As Dr. Natalie Alkiviadou, a senior fellow at The Future of Free Speech, puts it, “the absence of a clear and principled standard weakens the coherence of its hate speech jurisprudence.”¹¹⁵ Two reasons explain this. First, states enjoy a “margin of appreciation” that the Court has drawn unusually wide in hate speech cases, by reference to national cultural and historical particularities, even though it is generally restrictive about the margin it allows in freedom of expression cases.¹¹⁶ Second, its application of Article 17—the abuse clause, which removes expression from the Convention's protection altogether—has been inconsistent. Holocaust denial is categorically excluded on that basis, while the denial of the Armenian genocide is not. In *Perinçek v. Switzerland*, the Court reasoned that there was no comparable consensus on the facts. Alkiviadou calls this “an implicit hierarchy of historical suffering.”¹¹⁷

The consequence is “a very low threshold of hate being deemed prohibitable by the Court, such as ridicule or insult.”¹¹⁸ Alkiviadou also warns that instruments such as the German Network Enforcement Act (NetzDG), and now the DSA, “embolden private platforms to over-remove content and authoritarian regimes to replicate such models to suppress dissent.”¹¹⁹ There is no reason to expect this framework to be applied more carefully to AI-generated content than it has been to human speech.

113 Business and Human Rights Resource Centre, “France: Prosecutors Widen Criminal Investigation into X to Cover Grok-Generated Explicit Deepfakes,” January 2, 2026, <https://www.business-humanrights.org/en/latest-news/france-prosecutors-widen-criminal-investigation-into-x-to-cover-grok-generated-explicit-deepfakes/>.

114 Justin Hendrix, “Tracking the Paris Prosecutor's Investigation into Elon Musk's X,” *Tech Policy Press*, February 3, 2026, updated May 8, 2026, <https://www.techpolicy.press/tracking-the-paris-prosecutors-investigation-into-elon-musks-x/>.

115 Natalie Alkiviadou, *Hate Speech and the European Court of Human Rights* (Abingdon: Routledge, 2025), 205.

116 Alkiviadou, *Hate Speech and the European Court of Human Rights*, 11.

117 Alkiviadou, *Hate Speech and the European Court of Human Rights*, 2, 221.

118 Alkiviadou, *Hate Speech and the European Court of Human Rights*, 12.

119 Alkiviadou, *Hate Speech and the European Court of Human Rights*, 12.

The principal EU instrument dealing with hate speech is Council Framework Decision 2008/913/JHA, which requires member states to criminalize public incitement to violence or hatred on grounds of race, color, religion, descent, or national or ethnic origin. Member states may extend that protection to other characteristics; most have done so for sexual orientation and many for gender identity.¹²⁰ Since December 2021, the Commission has been trying to go further by adding hate speech and hate crime to the list of “EU crimes” in the Treaty on the Functioning of the European Union, which would let the EU set minimum criminal rules directly. That step requires unanimity in the Council, which the Commission has not obtained. Jacob Mchangama, Executive Director at The Future of Free Speech, and Alkiviadou have argued that the proposal fails the tests of legality, necessity, and proportionality.¹²¹

A person who uses a generative model to intentionally produce and then publish punishable hate speech would presumably be liable as though they had written it themselves.

The liability of providers is a harder question, but several routes are open. The DSA can apply to generative AI services, as evidenced by ChatGPT’s designation as a VLOSE, and providers of designated services face substantial fines for failing to address illegal content, including illegal hate speech. When a system generates new content rather than hosting someone else’s, it is highly unlikely to benefit from the safe harbors that shield hosting services.¹²² The Code of Conduct on Countering Illegal Hate Speech Online+, which became part of the DSA framework in January 2025, establishes an additional layer, since signatories commit to reviewing at least half of the notices they receive from designated monitoring reporters within twenty-four hours.¹²³ Such a short deadline incentivizes precautionary removal.

The AI Act does not address hate speech expressly. But, as section 4 explained, it requires

120 Ionel Zamfir and Colin Murphy, “Hate Speech and Hate Crime Targeting LGBTI People,” EPRS Briefing PE 767.219 (January 28, 2025): 1, https://www.europarl.europa.eu/RegData/etudes/BRIE/2025/767219/EPRS_BRI%282025%29767219_EN.pdf.

121 Jacob Mchangama and Natalie Alkiviadou, “A Cure Worse Than the Disease: Why the EU’s Proposal to Criminalize Hate Speech Violates Freedom of Expression — and Won’t Reduce Intolerance or Extremism,” *Journal of Free Speech Law* 6 (2026): 1183, 1187–88, <https://www.journaloffreespeechlaw.org/mchangamaalkiviadou2.pdf>.

122 Beatriz Botero Arcila, “Is It a Platform? Is It a Search Engine? It’s ChatGPT! The European Liability Regime for Large Language Models,” *Journal of Free Speech Law* 3 (2023): 460, <https://www.journaloffreespeechlaw.org/boteroarcila.pdf>.

123 European Commission, “Code of Conduct on Countering Illegal Hate Speech Online +,” integrated into the DSA framework on January 20, 2025, <https://digital-strategy.ec.europa.eu/en/library/code-conduct-countering-illegal-hate-speech-online>; European Commission, “First Results Published under the Revised Code of Conduct on Countering Illegal Hate Speech Online +,” April 10, 2026, <https://digital-strategy.ec.europa.eu/en/news/first-results-published-under-revised-code-conduct-countering-illegal-hate-speech-online>.

providers of GPAI models with systemic risk to assess and mitigate negative effects on fundamental rights and on “the society as a whole,” and the GPAI Code of Practice identifies “hateful” and “radicalizing” content among the risks to be mitigated.¹²⁴ The Code is not binding, but it is endorsed by most major providers, and it is the document against which the AI Office will measure them. It is reasonable to expect the Commission to treat hate speech—and certainly illegal hate speech—as a systemic risk capable of triggering enforcement.

¹²⁴ European Commission, “Safety and Security,” in *General Purpose AI Code of Practice*, 34.

6. DEFAMATION

What happens when an AI model generates defamatory content? When it states, falsely, that someone has committed fraud or been convicted of a crime? AI is testing European defamation law, and the difficulty is fitting AI-generated content into a framework built around the distinction between primary publishers, intermediaries, and republishers of defamatory material.

There is no single EU defamation law. Each member state applies its own framework, shaped by national history and by its own reading of how reputation should be balanced against expression under the Convention and the Charter. That balance is struck restrictively in a significant part of the EU. Many member states still criminalize defamation, some with penalties of imprisonment,¹²⁵ and the Council of Europe has for years urged them to repeal those provisions, warning that criminal defamation law chills public debate.¹²⁶

The Strasbourg Court has been an unreliable guardian of free expression. The Court polices disproportionate sanctions,¹²⁷ but its case law shows a growing readiness to accept restrictions on online expression. Notably, the Court has stated that the risk of harm posed by content online is “certainly higher” than that posed by the press,¹²⁸ and it found no violation of Article 10 when Estonia held a news portal responsible for defamatory comments left by its readers,¹²⁹ or when France convicted a politician for failing to delete third-party comments from his Facebook wall (section 2).¹³⁰ Scholars have described this as an increase in “the consideration given to restrictions to free speech” as the internet has developed.¹³¹

125 Mario Viola de Azevedo Cunha and Luc Steinberg, eds., “Decriminalisation of Defamation Factsheet,” Centre for Media Pluralism and Media Freedom, European University Institute, January 2019, https://cmpf.eui.eu/wp-content/uploads/2019/01/decriminalisation-of-defamation_Infographic.pdf.

126 Council of Europe, “Freedom of Expression: Defamation,” <https://www.coe.int/en/web/freedom-expression/defamation>.

127 *Cumpănă and Mazăre v. Romania* [GC], App. No. 33348/96, Eur. Ct. H.R. (December 17, 2004), para. 115, <https://hudoc.echr.coe.int/eng?i=001-67816>.

128 *Editorial Board of Pravoye Delo and Shtetel v. Ukraine*, App. No. 33014/05, Eur. Ct. H.R. (May 5, 2011, final August 5, 2011), <https://hudoc.echr.coe.int/eng?i=001-104685>.

129 *Delfi AS v. Estonia* [GC], App. No. 64569/09, Eur. Ct. H.R. (June 16, 2015), <https://hudoc.echr.coe.int/eng?i=001-155105>.

130 *Sanchez v. France*, App. No. 45581/15.

131 Oreste Pollicino and Marco Bassini, “Free Speech, Defamation and the Limits to Freedom of Expression in the EU: A Comparative Analysis,” in *Research Handbook on EU Internet Law*, ed. Andrej Savin and Jan Trzaskowski (Cheltenham: Edward Elgar, 2014), 530.

Recent decisions by German courts provide some hints on how the issue of AI-generated defamation may be approached in the EU, or at least in Germany.

The first decision, issued by the Regional Court of Frankfurt am Main in September 2025, rejected an application by a group of plastic surgeons concerning an allegedly false Google AI Overview about medical procedures. It found that the content was not false when read in full context, but it clarified that Google may be held responsible for incorrect AI-generated statements.¹³²

The second decision went further. In May 2026, the Regional Court of Munich I granted an injunction against Google over AI Overviews that told users that two Munich publishing houses were known for questionable business practices and were associated with scams. The court held that Google was directly responsible, not merely as an intermediary, because an AI Overview is the defendant's own statement, generated by its own system. The disputed accusations appeared in none of the sources the overview linked to. The court rejected the argument that users can and should verify AI output for themselves, reasoning that the statement was self-contained and intelligible on its face, so a reader had no reason to consult the sources. It prohibited the specific statements.¹³³

In the Munich case, the statements did not come from third-party content, and the claimants had asked Google to remove them. AI systems that generate new content, rather than merely retrieving it as a search engine does, are highly unlikely to benefit from the "safe harbors" exempting them from liability.¹³⁴ In the past, the German Federal Court of Justice ruled that a search engine provider could be held liable for failing to take reasonable steps to prevent defamatory suggestions in its autocomplete function.¹³⁵ General terms and conditions and disclaimers stating that the AI service may provide false or incorrect information are also unlikely to protect AI providers from liability.¹³⁶

132 Katharina Wagner, "Germany: Court Holds Google Liable for Incorrect AI Overviews," *Global Legal Monitor*, Law Library of Congress, July 17, 2026, <https://www.loc.gov/item/global-legal-monitor/2026-07-17/germany-court-holds-google-liable-for-incorrect-ai-overviews/>.

133 Landgericht München I, 26 O 869/26; see Giovanni Vetrugno, "Who Speaks When the Algorithm Speaks? A German Ruling on AI Overviews," *Oxford Business Law Blog*, July 16, 2026, <https://blogs.law.ox.ac.uk/oblb/blog-post/2026/07/who-speaks-when-algorithm-speaks-german-ruling-ai-overviews>.

134 Botero Arcila, "Is It a Platform?," 460.

135 Claudio Novelli, Federico Casolari, Philipp Hacker, Giorgio Spedicato, and Luciano Floridi, "Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity," *Computer Law & Security Review* 55 (2024): 7, <https://doi.org/10.1016/j.clsr.2024.106066>.

136 Graziana Kastl-Riemann, "Regulation of Generative AI Speech: An EU Perspective," in *The Oxford Handbook of the Foundations and Regulation of Generative AI*, ed. Philipp Hacker et al. (Oxford University Press, 2025), <https://doi.org/10.1093/oxford-hb/9780198940272.013.0022>.

It is understandable that AI companies cannot benefit from safe harbors when they provide content they themselves have autonomously generated. At the same time, courts and policymakers should be cautious about imposing overly strict obligations for defamation. In 2024, users discovered that ChatGPT would abruptly end conversations when certain names appeared, apparently because the provider had hard-coded blocks around individuals who had threatened to sue, among them Brian Hood, an Australian mayor who had complained about false statements the system made about him.¹³⁷

The potential liability of the user generating the content is also relevant. Someone who prompts a chatbot and then publishes its answer becomes its publisher. In many member states, including France¹³⁸ and Germany,¹³⁹ repeating a defamatory allegation is treated as equivalent to making it. Users may still rely on the ordinary defenses, such as showing the allegation had a sufficient factual basis or that they acted in good faith.¹⁴⁰

European law offers a second route where the false statement concerns an identifiable person: data protection. Inaccurate personal data engages the GDPR, which gives the individual rights to rectification and erasure and provides a right to compensation for material or nonmaterial damage. However, compensation has significant limitations. The Luxembourg Court has held that an infringement of the GDPR is not in itself enough to found a claim, and the person must also prove actual damage.¹⁴¹ Hence, rectification is the realistic remedy in most cases. The privacy organization noyb has pursued this route on behalf of individuals about whom ChatGPT generated false and damaging statements, including a Norwegian man the system described as a convicted murderer.¹⁴²

137 Mike Masnick, "The Curious Case of ChatGPT's Banned Names: Hard-Coding Blocks to Avoid Nuisance Threats," *Techdirt*, December 3, 2024, <https://www.techdirt.com/2024/12/03/the-curious-case-of-chatgpts-banned-names-hard-coding-blocks-to-avoid-nuisance-threats>; Benj Edwards, "Certain Names Make ChatGPT Grind to a Halt, and We Know Why," *Ars Technica*, December 2, 2024, <https://arstechnica.com/information-technology/2024/12/certain-names-make-chatgpt-grind-to-a-halt-and-we-know-why>.

138 Dalloz, "Fiches d'orientation: Diffamation," September 2022, <https://www.dalloz.fr/documentation/Document?id=DZ/OASIS/000328> (Fr.).

139 Strafgesetzbuch [StGB] [Penal Code] §§ 186–87 (Ger.), https://www.gesetze-im-internet.de/englisch_stgb/englisch_stgb.html.

140 ECtHR, *Guide on Article 10*, paras. 232–37.

141 Regulation (EU) 2016/679 (General Data Protection Regulation) arts. 16, 17 and 82; Case C-300/21, *UI v. Österreichische Post AG*, judgment of May 4, 2023, ECLI:EU:C:2023:370.

142 noyb, "AI Hallucinations: ChatGPT Created a Fake Child Murderer," March 20, 2025, <https://noyb.eu/en/ai-hallucinations-chatgpt-created-fake-child-murderer>; Natasha Lomas, "ChatGPT Hit with Privacy Complaint over Defamatory Hallucinations," *TechCrunch*, March 19, 2025, <https://techcrunch.com/2025/03/19/chatgpt-hit-with-privacy-complaint-over-defamatory-hallucinations>.

7. PROTECTION OF MINORS

7.1 Background

Child protection online was one of the most active discussion areas of EU digital policy in the period covered by this chapter. The importance of the considerations at stake and the tension between them make it a particularly challenging topic. The protection of minors is an aim no one disputes. Still, the mechanisms chosen—age verification, access restrictions, design mandates—can reach everyone, including adults, whose access to lawful information and whose privacy may both be curtailed. Minors are also rights holders. The Committee on the Rights of the Child recognizes that children’s freedom of expression includes seeking, receiving, and imparting information, and has pointed out that restrictions in the digital space, including filters and safety measures, must be lawful, necessary, and proportionate.¹⁴³

7.2 CSAM

In the EU, CSAM is outlawed under Directive 2011/93/EU, which covers “realistic images of a child” within its definition, language broad enough to reach computer-generated material. Member states retain discretion not to criminalize the production and possession of such realistic images where the producer keeps them solely for private use, no material depicting a real child was used to make them, and there is no risk of dissemination.¹⁴⁴ The AI Act now adds a prohibition on AI systems that generate such material, as discussed in section 8.

The directive is being revised to take AI into account. In June 2026, the Parliament and the Council reached a provisional agreement on a recast that would make it a criminal offense to design or adapt an AI system to produce CSAM, or to distribute such a system. The recast

¹⁴³ UN Committee on the Rights of the Child, General Comment No. 25 (2021) on Children’s Rights in Relation to the Digital Environment, CRC/C/GC/25, March 2, 2021, paras. 58 and 59. See also Convention on the Rights of the Child, November 20, 1989, 1577 U.N.T.S. 3, art. 13.

¹⁴⁴ Directive 2011/93/EU on Combating the Sexual Abuse and Sexual Exploitation of Children and Child Pornography, 2011 O.J. (L 335) 1, arts. 2(c)(iv), 5(7)–(8); Katalin Parti and Judit Szabó, “The Legal Challenges of Realistic and AI-Driven Child Sexual Abuse Material: Regulatory and Enforcement Perspectives in Europe,” *Laws* 13, no. 6 (2024): art. 67, 11, <https://doi.org/10.3390/laws13060067>.

would also criminalize possessing and distributing instructions on how to produce that material, “for instance through a book, an online guide or a prompt.”¹⁴⁵ At the time of writing, the directive has not yet been adopted.

The Strasbourg Court has left little room for free expression arguments in this domain. In *Karttunen v. Finland*, an artist seeking to provoke debate about the accessibility of CSAM was convicted after exhibiting, in a gallery show, images downloaded from the internet; the domestic court waived the sanction and ordered the images confiscated. The Court accepted that the conviction interfered with her freedom of expression, but held that the reasons given by the domestic courts were relevant and sufficient and the interference proportionate, and declared the complaint manifestly ill-founded.¹⁴⁶

The main AI-specific enforcement action at EU level concerns X. EU institutions are also negotiating the proposed Child Sexual Abuse Regulation, which would replace the temporary derogation that allows providers to scan communications voluntarily. The regulation has not been adopted, and negotiations have been particularly challenging, owing to concerns that it would amount to generalized monitoring of private communications. Sections 8 and 10 discuss these issues.

7.3 Age-Based Restrictions

Much of the EU’s activity, however, has not been about AI-generated material. Its policy initiatives have primarily dealt with keeping children away from certain services. The Commission’s guidelines on the protection of minors under the DSA, published in July 2025, recommend age verification for platforms distributing pornography or gambling, and for services subject to a legal minimum age under national rules, including defined categories of social media. They recommend less intrusive age estimation in other cases, including where a platform’s own terms set a minimum age below eighteen because of identified risks.¹⁴⁷ As the DSA increasingly applies to AI products and services, these

¹⁴⁵ Council of the European Union, “Combatting Child Sexual Abuse: EU Agrees Stronger Criminal Law Rules and Enhanced Support to Victims,” press release, June 22, 2026, <https://www.consilium.europa.eu/en/press/press-releases/2026/06/22/combatting-child-sexual-abuse-eu-agrees-stronger-criminal-law-rules-and-enhanced-support-to-victims/>.

¹⁴⁶ *Karttunen v. Finland* (dec.), App. No. 1685/10, Eur. Ct. H.R. (May 10, 2011), <https://hudoc.echr.coe.int/eng?i=001-104816>; see Rónán Ó Fathaigh, “*Karttunen v. Finland*: Child Pornography and Freedom of Expression,” *Strasbourg Observers*, October 10, 2011, <https://strasbourgobservers.com/2011/10/10/child-pornography-and-freedom-of-expression>.

¹⁴⁷ European Commission, “Guidelines on the Protection of Minors under Regulation (EU) 2022/2065,” July 14, 2025, <https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-protection-minors>.

recommendations matter for them too.

Importantly, the Commission sponsored the development of a white-label, open-source age verification blueprint, and in April 2026 issued a recommendation establishing a common EU framework, asking member states to draw up implementation plans and make a solution available by December 31, 2026.¹⁴⁸ Despite the Commission's privacy assurances, a security researcher reported bypassing the prototype's protections in under two minutes shortly after it was presented.¹⁴⁹

These initiatives have been accompanied by enforcement actions. In 2025 and 2026, the Commission issued preliminary findings under the DSA against four adult content platforms, against Meta over access to Facebook and Instagram by children under thirteen, and against TikTok over addictive design and the safety of minors' accounts. It has also opened proceedings against Snapchat.¹⁵⁰

The political focus on measures related to the protection of minors has been considerable. In October 2025, twenty-five of the twenty-seven member states signed the Jutland Declaration supporting a "digital age of majority"; Estonia and Belgium did not.¹⁵¹ In November 2025, the European Parliament issued a resolution calling for a harmonized European digital minimum age of sixteen for social media, video-sharing platforms, and AI companions, with thirteen as an absolute floor and parental consent between the two.¹⁵² In July 2026, a special panel of experts convened by the Commission recommended a harmonized EU-wide restriction on access for children under thirteen, and defined the services covered as "social media+," a category that expressly includes AI chatbots.¹⁵³ By

148 Commission Recommendation (EU) 2026/1035 of 29 April 2026 establishing a common approach to EU-wide age verification technologies, <https://digital-strategy.ec.europa.eu/en/library/commission-sets-out-common-approach-eu-wide-age-verification-technologies>.

149 "EU age verification app can be hacked in 2 minutes, claims security expert," *Cybernews*, April 16, 2026, <https://cybernews.com/security/eu-age-verification-app-hack/>; "EU's 'Safe by Design' Age-Verification App Cracked in Minutes, Raising Data Security Fears," *MediaNama*, April 2026, <https://www.medianama.com/2026/04/223-eu-age-verification-app-hacked-data-security-fears/>.

150 See European Commission, Digital Services Act enforcement actions, <https://digital-strategy.ec.europa.eu/en/policies/dsa-enforcement>. Preliminary findings were issued against four adult content platforms and proceedings opened against Snapchat in March 2026; against Meta over under-13 access in April 2026; and against TikTok over addictive design in February 2026 and over the safety of minors' accounts in July 2026.

151 Jessica Galissaire, Salma Gomez, and Lena-Maria Böswald, "Social Media Age Gating in the EU," interface, May 19, 2026, <https://www.interface-eu.org/publications/social-media-age-gating-eu> (recording that Belgium and Estonia did not sign, and that the Belgian position followed a veto by the Flemish media minister). Sources differ on the exact day of adoption in October 2025.

152 European Parliament resolution of 26 November 2025 on the protection of minors online, 2025/2060(INI), para. 28.

153 European Commission, Special Panel on Child Safety Online, *Child Safety Online: Protecting and Empowering Minors in a Digital World* (July 13, 2026), https://commission.europa.eu/document/download/4c5d55f1-9072-4bb9-a263-094461525ca6_en.

May 2026, twenty-three of the twenty-seven member states were at least contemplating legislation to restrict or ban access to social media under a certain age.¹⁵⁴

Shortly before this chapter was published, the Commission proposed a new regulation, the EU KIDS Act.¹⁵⁵ It would prohibit access to social media for children under thirteen and set an EU-wide minimum age of fifteen for opening a full account. Children between thirteen and fifteen would be able to use accounts with limited features. It would cover social media, video-sharing platforms, and online games, and would require providers to disable AI chatbots and companions by default for minors and to prevent interactions that create emotional dependency. Providers would have to show that their services are “age-appropriate and safe by design,” and to verify age through the EU age verification app or other means.¹⁵⁶

Restrictions aimed at conversational AI are also starting to appear. An Italian law conditions access to AI technologies by children under fourteen on parental consent.¹⁵⁷ The Italian Data Protection Authority fined a developer of the companion app Replika €5 million in April 2025, finding that it had no age verification mechanism at all despite claiming that minors were not allowed on the platform.¹⁵⁸ In July 2026, the authority fined Character Technologies €158,000 and ordered it to put in place effective age checks, a cooling-off period to stop blocked minors from simply reregistering, and private-by-default profiles for minors.¹⁵⁹

7.4 Other Initiatives

At EU level, the DSA’s obligations for the protection of minors in Article 28 apply to online platforms. Because ChatGPT was designated as a VLOSE rather than an online platform, it

154 Galissaire, Gomez, and Böswald, “Social Media Age Gating in the EU.”

155 European Commission, “EU KIDS Act to restrict social media platforms’ access to children in the EU,” press release IP/26/1890, Brussels, September 17, 2026, https://ec.europa.eu/commission/presscorner/detail/en/ip_26_1890.

156 European Commission, “EU KIDS Act to restrict social media platforms’ access to children in the EU.”

157 Legge 23 settembre 2025, n. 132 [Law No. 132 of September 23, 2025], in force October 10, 2025 (It.).

158 Garante per la protezione dei dati personali, provvedimento of April 10, 2025 [doc. web no. 10130115], <https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/10130115>; see European Data Protection Board, “AI: The Italian Supervisory Authority Fines Company Behind Chatbot ‘Replika,’” May 21, 2025, https://www.edpb.europa.eu/news/ai-the-italian-supervisory-authority-fines-company-behind-chatbot-replika_en.

159 Garante per la protezione dei dati personali, “Intelligenza artificiale: il Garante privacy sanziona Character.AI. Rilevate criticità anche nella tutela dei minori e nei sistemi di verifica dell’età” [Artificial intelligence: the privacy Garante sanctions Character.AI. Shortcomings also found in the protection of minors and in age verification systems], press release, July 9, 2026, <https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/10269594>.

is unclear whether Article 28 applies to it at all.¹⁶⁰ The Parliament adopted a resolution that calls for “a prohibition of human-like features” in chatbots aimed at minors.¹⁶¹

¹⁶⁰ European Commission, “Commission Designates ChatGPT.”

¹⁶¹ European Parliament resolution of 26 November 2025 on the protection of minors online, para. 77.

8. EXPLICIT CONTENT

The EU has been particularly active on NCII. Much of this activity was precipitated by the rollout of image editing in Grok in late December 2025. Users employed it to digitally undress real people in ordinary photographs on request, and to post the results on X. Between December 25, 2025, and January 1, 2026, a European research organization collected 50,000 mentions of the Grok account from 8,500 unique users, along with 20,000 images the account generated. It found that 53 percent of the images contained individuals in minimal attire, 81 percent of whom presented as women.¹⁶² Other organizations put the total higher, at around three million sexualized images in eleven days.¹⁶³

The scandal resulted in an extensive regulatory response. The Commission ordered X to preserve internal documents relating to Grok and then opened formal proceedings under the DSA, examining whether X had assessed the risks of deploying Grok before launch and whether its mitigation was adequate.¹⁶⁴ French investigators searched X's Paris offices as part of a criminal investigation that now encompasses both the sexual imagery and the Holocaust-denial output discussed in section 5. In February 2026, the Irish Data Protection Commission opened an inquiry under the Data Protection Act 2018 in relation to "the apparent creation, and publication on the X platform, of potentially harmful, non-consensual intimate and/or sexualised images," including of children.¹⁶⁵ In March 2026, the Amsterdam District Court ordered X to stop generating nonconsensual sexualized imagery and CSAM in the Netherlands.¹⁶⁶

¹⁶² AI Forensics, *Grok Unleashed: Grok Generating Flood of Sexualized Images of Women, Including Minors, and Extremist Propaganda* (January 2, 2026), <https://aiforensics.org/uploads/GrokUnleashed.pdf>.

¹⁶³ Center for Countering Digital Hate, *Grok Floods X with Sexualized Images of Women and Children* (January 22, 2026), <https://counter-hate.com/research/grok-floods-x-with-sexualized-images/>.

¹⁶⁴ European Commission, "Commission Investigates Grok and X's Recommender Systems."

¹⁶⁵ Data Protection Commission (Ireland), "Data Protection Commission Opens Investigation into X (XIUC)," February 17, 2026, <https://www.dataprotection.ie/en/news-media/press-releases/data-protection-commission-opens-investigation-x-xiuc>.

¹⁶⁶ Rechtbank Amsterdam, ECLI:NL:RBAMS:2026:3106, judgment of March 26, 2026; see Ramsha Jahangir, "Dutch Court Orders X, Grok to Stop AI-Generated Sexual Abuse Content," *Tech Policy Press*, March 26, 2026, <https://www.techpolicy.press/dutch-court-orders-x-grok-to-stop-ai-generated-sexual-abuse-content/>.

X's own response was incremental. On January 9, 2026, it limited image generation on the platform to paying subscribers,¹⁶⁷ and in the following days blocked requests to undress real people in images.

The legislative response was a new prohibition in the AI Act, added by the AI Omnibus and applicable from December 2, 2026. This new provision bans AI systems that generate or manipulate CSAM or NCII. For providers, as section 3 explained, it reaches any system where such output is a reasonably foreseeable and reproducible outcome achievable without significant technical modification, and where the system lacks adequate technical safeguards. It applies to deployers only where they use a system intentionally for that purpose.¹⁶⁸

The prohibition on NCII applies only where the depicted person has not provided “freely-given, specific, informed, unambiguous and explicit consent for that generation or manipulation.” The CSAM limb has no equivalent. Protecting consensual intimate content between adults matters, especially since that outcome was not guaranteed. Indeed, six Dutch bodies had jointly called for the European ban to go further. Their position was that “for legislation to be effective, it is necessary to ban these tools completely, even in those cases where the depicted individuals may have consented to the creation of those images,” because current law “makes it primarily possible to take action against individual offenders” and offers “insufficient options for taking action against the tools themselves.”¹⁶⁹

The prohibition in the AI Act is complemented by criminal rules. Directive (EU) 2024/1385 on combating violence against women and domestic violence requires member states to criminalize the nonconsensual production, manipulation, and sharing of intimate material, and states expressly that this should cover the fabrication of deepfakes where the material appreciably resembles an existing person and would falsely appear authentic.¹⁷⁰

Finally, the DSA requires VLOPs and VLOSEs to assess and mitigate systemic risks, and

¹⁶⁷ Ram Iyer, “X Restricts Grok’s Image Generation to Paying Subscribers Only After Drawing the World’s Ire,” *TechCrunch*, January 9, 2026, <https://techcrunch.com/2026/01/09/x-restricts-groks-image-generation-to-paying-subscribers-only-after-drawing-the-worlds-ire/>.

¹⁶⁸ Regulation (EU) 2026/1744, amending Artificial Intelligence Act art. 5; see section 3 above.

¹⁶⁹ Autoriteit Consument en Markt, “Dutch Regulators, the Dutch Police, and the Dutch Public Prosecution Service Welcome a European Ban on ‘AI Nudify Apps and Websites,’” April 1, 2026, <https://www.acm.nl/en/publications/dutch-regulators-dutch-police-and-dutch-public-prosecution-service-welcome-european-ban-ai-nudify-apps-and-websites>.

¹⁷⁰ Directive (EU) 2024/1385 of 14 May 2024 on combating violence against women and domestic violence, 2024 O.J. (L, 2024/1385), in force June 13, 2024.

states that services used principally to disseminate pornographic content should process notices and remove nonconsensually shared intimate or manipulated material without delay.¹⁷¹ The GDPR also provides a route to deal with NCII, because intimate images are personal data, but its practical usefulness against gender-based abuse has been limited.¹⁷²

¹⁷¹ Digital Services Act arts. 34–35 and recital 87.

¹⁷² Ionel Zamfir and Colin Murphy, “Cyberviolence Against Women in the EU,” EPRS Briefing (2024): 6.

9. COPYRIGHT

Two issues are particularly relevant to copyright and generative AI. The first concerns the inputs—specifically, the data—used to develop the model. The second concerns the outputs—the content—generated by the model. How the law treats outputs can affect the information and perspectives people can access from these systems. Notably, liability for outputs gives providers incentives to filter content generation, and filters built to avoid infringement can catch quotation, criticism, and parody, which EU law expressly protects.¹⁷³

Training a generative AI model requires very large datasets, and those datasets draw heavily on material that is freely accessible online. Their composition is typically treated as confidential and commercially sensitive. A number of lawsuits allege that datasets contain copyright-protected works.¹⁷⁴ To assemble them, developers use web crawlers, that is, software that browses and copies at scale.¹⁷⁵

Copyright gives authors exclusive rights, including rights of reproduction and communication to the public.¹⁷⁶ The Copyright in the Digital Single Market Directive (CDSM Directive) created two exceptions for text and data mining (TDM). Article 3 permits research organizations and cultural heritage institutions to mine lawfully accessible content for scientific research. Rights holders cannot opt out in this case. Article 4 permits anyone, including commercial developers, to mine lawfully accessible content for any purpose, but mining is possible only if the rights holder has not expressly reserved the use.¹⁷⁷

Neither provision was drafted with generative models in mind, and there has been much debate over whether the TDM exceptions protect the training of AI models. A number of

¹⁷³ Caroline De Cock, “From Upload to Output Filters? How the AI Act’s Code of Practice Could Threaten Freedom of Expression and Thought,” *Tech Policy Press*, January 17, 2025, <https://www.techpolicy.press/from-upload-to-output-filters-how-the-ai-acts-code-of-practice-could-threaten-freedom-of-expression-and-thought/>.

¹⁷⁴ Tristan Marcellin and Filippo Cassetti, “AI and Copyright: The Training of General-Purpose AI,” EPRS Note PE 769.585 (April 28, 2025), <https://epthinktank.eu/2025/04/28/ai-and-copyright-the-training-of-general%E2%80%91purpose-ai>.

¹⁷⁵ Marcellin and Cassetti, “AI and Copyright.”

¹⁷⁶ Directive 2001/29/EC on the Harmonisation of Certain Aspects of Copyright and Related Rights in the Information Society, 2001 O.J. (L 167) 10, arts. 2–3.

¹⁷⁷ Directive 2019/790 on Copyright and Related Rights in the Digital Single Market, 2019 O.J. (L 130) 92, arts. 3–4.

scholars have argued that Article 4 does reach it, subject to the rights reservation.¹⁷⁸ Others argue that text and data mining is not synonymous with AI training and that the exceptions do not cover all the actions required for models to be trained.¹⁷⁹

National courts have begun to grapple with this question. In Germany, the Higher Regional Court of Hamburg ruled that downloading and analyzing images to build and validate a training dataset falls within the TDM exceptions; that a noncommercial research organization can rely on the research exception even where commercial developers use the dataset it produces, provided private companies do not control it or enjoy preferential access; and that a rights holder's reservation is effective only if a machine can act on it.¹⁸⁰

Other courts have reached different conclusions. In *GEMA v. OpenAI*, the Regional Court of Munich I held that song lyrics retained in a model's parameters are reproduced within the meaning of German copyright law, and that the TDM exceptions do not cover that retention.¹⁸¹

On outputs, the court found that regurgitation of lyrics in AI outputs engages the rights of reproduction and communication to the public. It also held that the provider, not the user, was liable, because the material could be extracted with simple prompts.¹⁸²

In a different case, the same court held that TDM exceptions did not apply, in part because the training material had been obtained by circumventing YouTube's technical protection measures.¹⁸³

At the time of writing, the Luxembourg Court has not ruled on this topic. In *Like Company v. Google Ireland*, the Court has been asked whether a chatbot's output reproducing

¹⁷⁸ See Martin Senftleben, "Generative AI and Author Remuneration," *IIC — International Review of Intellectual Property and Competition Law* 54, no. 10 (2023): 1535–60, <https://doi.org/10.1007/s40319-023-01399-4>.

¹⁷⁹ See Eleonora Rosati, "Is Text and Data Mining Synonymous with AI Training?," *Journal of Intellectual Property Law & Practice* 19, no. 12 (2024): 851–52, <https://doi.org/10.1093/jiplp/jpae092>; Tim W. Dornis and Sebastian Stober, Copyright Law and Generative AI Training: Technological and Legal Foundations [Urheberrecht und Training generativer KI-Modelle: Technologische und juristische Grundlagen] (Baden-Baden: Nomos, 2024), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4946214.

¹⁸⁰ Oberlandesgericht Hamburg, 5 U 104/24, judgment of December 10, 2025 (*Kneschke v. LAION*) (Ger.); see "Machine-Readable Opt-Outs and AI Training: Hamburg Court Clarifies Copyright Exceptions," Norton Rose Fulbright, December 2025, <https://www.insidetechnology.com/blog/2025/12/machine-readable-opt-outs-and-ai-training-hamburg-court-clarifies-copyright-exceptions>. A further appeal to the Federal Court of Justice was allowed.

¹⁸¹ Landgericht München I, 42 O 14139/24 (*GEMA v. OpenAI*).

¹⁸² Landgericht München I, 42 O 14139/24 (*GEMA v. OpenAI*).

¹⁸³ Landgericht München I, 42 O 763/25, judgment of July 31, 2026 (*GEMA v. Suno*) (Ger.); see European IP Helpdesk, "German Court Finds Suno Liable for Copyright Infringement over AI-Generated Music," August 28, 2026, https://intellectual-property-helpdesk.ec.europa.eu/news-events/news/german-court-finds-suno-liable-copyright-infringement-over-ai-generated-music-2026-08-28_en.

press content is a communication to the public, whether training a model is an act of reproduction, and whether Article 4 covers it.¹⁸⁴

Policymakers are considering which approach to adopt. In March 2026, the European Parliament adopted a resolution on copyright and generative AI, which calls for rules clarifying licensing and for voluntary collective licensing. It also asks the Commission to examine a possible mechanism for “immediate, fair and proportionate remuneration for past uses” of protected works by providers of GPAI models and systems where licensing markets do not yet exist. The resolution further states that rights holders, especially in the press and news media, must have “full control over the digital use of their content by AI systems and models for training purposes,” and that uses beyond training, such as inferencing and retrieval-augmented generation, may take place “only with the express consent of the rights holders.”¹⁸⁵

In parallel, the Commission has conducted consultations on the review of the CDSM Directive and on a targeted initiative for a “copyright environment for European creativity and innovation.”¹⁸⁶

The AI Act includes two relevant obligations. Providers of GPAI models must put in place a policy to comply with EU copyright law, including the Article 4 reservation mechanism, and must publish a sufficiently detailed summary of the content used for training, following a template issued by the Commission.¹⁸⁷

As for the protection of AI-generated content, EU law requires an “original work” that is the author’s own intellectual creation, a standard understood to require human creativity.¹⁸⁸ Where a person uses AI as a tool and contributes the creative choices, the result is likely protected.¹⁸⁹ Where the human contribution is minimal, it is likely not.¹⁹⁰

184 Like Company, C-250/25.

185 European Parliament resolution of 10 March 2026 on copyright and generative artificial intelligence – opportunities and challenges, P10_TA(2026)0066, 2025/2058(INI), paras. 2, 5–6, 9, 21.

186 European Commission, call for evidence on the review of Directive (EU) 2019/790 and on a targeted initiative for a better copyright environment for European creativity and innovation, launched May 13, 2026, closed June 25, 2026.

187 Artificial Intelligence Act art. 53(1)(c)–(d).

188 Novelli et al., “Generative AI in EU Law,” 11.

189 Novelli et al., “Generative AI in EU Law,” 11.

190 Novelli et al., “Generative AI in EU Law,” 11.

10. AI-POWERED SURVEILLANCE

Surveillance directly affects freedom of expression. It discourages public participation and criticism of powerful people and institutions. The AI Act declares that real-time remote biometric identification in public spaces—one of many ways of deploying mass surveillance—“may affect the private life of a large part of the population, evoke a feeling of constant surveillance and indirectly dissuade the exercise of the freedom of assembly and other fundamental rights.”¹⁹¹

In *Glukhin v. Russia*, the police used facial-recognition technology to identify a man who had staged a solo demonstration in the Moscow metro, drawing on images of the protest circulating online and on footage from metro cameras. The Strasbourg Court found violations of the rights to privacy and freedom of expression, highlighting the chilling effect of these measures on political participation. It described facial recognition as highly intrusive, though it did not rule on whether the technology as such is compatible with the Convention.¹⁹²

Mass surveillance has become significantly easier with AI, which reduces the personnel and time required. The most advanced deployments of mass surveillance are in authoritarian states, and China both operates¹⁹³ and exports such systems.¹⁹⁴ But the EU has its own examples. In March 2025, the Hungarian Parliament passed legislation banning Pride events and authorizing the authorities to use facial recognition to identify organizers and

¹⁹¹ Artificial Intelligence Act, recital 32.

¹⁹² *Glukhin v. Russia*, App. No. 11519/20, Eur. Ct. H.R. (July 4, 2023); see Cristina Cocito, “*Glukhin v. Russia*: facial recognition considered highly intrusive but not inconsistent with fundamental rights,” *Strasbourg Observers*, January 9, 2024, <https://strasbourgobservers.com/2024/01/09/glukhin-v-russia-facial-recognition-is-a-highly-intrusive-technology-but-the-court-abstains-from-considering-a-potential-inconsistency-with-fundamental-rights/>.

¹⁹³ Jessie Yeung, “China’s Censorship and Surveillance Were Already Intense. AI Is Turbocharging Those Systems,” *CNN*, December 3, 2025, <https://www.cnn.com/2025/12/04/china/china-ai-censorship-surveillance-report-intl-hnk>.

¹⁹⁴ Dahlia Peterson, *Designing Alternatives to China’s Repressive Surveillance State*, CSET Policy Brief (Washington, DC: Center for Security and Emerging Technology, October 2020), <https://cset.georgetown.edu/wp-content/uploads/CSET-Designing-Alternatives-to-Chinas-Surveillance-State.pdf>.

attendees.¹⁹⁵ Around 100,000 people marched in Budapest that June in defiance of the ban.

France provides another example. Ahead of the 2024 Olympic Games, it authorized algorithmic video surveillance in public spaces, permitting AI systems to scan crowds in real time for defined events.¹⁹⁶ After the Games, the Paris police prefect publicly supported making the power permanent, and the then prime minister undertook to generalize the methods used during the Olympics.¹⁹⁷

The AI Act responds to these challenges in part. The Act does not ban mass surveillance as such, but it prohibits several of the practices that make it possible. It bans placing on the market or using AI systems that create or expand facial recognition databases through the untargeted scraping of images from the internet or CCTV footage. It also bans social scoring where the score leads to detrimental treatment in unrelated contexts, or treatment disproportionate to the behavior. In addition, it bans predictive policing that assesses criminal risk solely on the basis of profiling or personality traits. It also bans biometric categorization systems that infer race, political opinions, trade union membership, religious or philosophical beliefs, sex life, or sexual orientation. Finally, it imposes a near-total prohibition on real-time remote biometric identification in publicly accessible spaces for law enforcement, permitting it only for a targeted search for victims or missing persons, the prevention of a specific and imminent threat to life or of a terrorist attack, or the localization of a suspect in specified serious offenses. In those cases, it requires an authorization from a judicial or independent administrative authority, a fundamental rights impact assessment, and strict limits of time and place.¹⁹⁸

However, the limits of the protection offered by the AI Act are substantial. First, retrospective identification is not prohibited. The analysis of footage after an event—“post-

195 Human Rights Watch, “Hungary Bans LGBT Pride Events,” March 20, 2025, <https://www.hrw.org/news/2025/03/20/hungary-bans-lgbt-pride-events>.

196 Library of Congress, *Global Legal Monitor*, “France: New Law Establishes Legal Framework for 2024 Olympic and Paralympic Games,” November 15, 2023, <https://www.loc.gov/item/global-legal-monitor/2023-11-15/france-new-law-establishes-legal-framework-for-2024-olympic-and-paralympic-games/>.

197 “Le préfet de police de Paris se dit favorable à une prolongation du recours à la vidéosurveillance algorithmique” [The Paris police prefect says he favors extending the use of algorithmic video surveillance], *Le Monde*, September 25, 2024, https://www.lemonde.fr/societe/article/2024/09/25/le-prefet-de-police-de-paris-se-dit-favorable-a-une-prolongation-du-recours-a-la-videosurveillance-algorithmique_6333125_3224.html (Fr.).

198 Artificial Intelligence Act art. 5(1)(c)–(h) and art. 5(2)–(7); see European Commission, Guidelines on Prohibited Artificial Intelligence Practices (February 2025).

remote” biometric identification—is classified as high-risk, but not prohibited. The chilling effect on assembly can happen whether the identification happens in real time or a week later.¹⁹⁹ Second, the real-time prohibition applies only to publicly accessible spaces. Third, the ban on biometric categorization includes a law-enforcement carve-out, and the ban on emotion recognition reaches only workplaces and educational institutions. Finally, the AI Omnibus postponed the high-risk obligations at issue here to December 2, 2027.

Crucially, the AI Act contains a fundamental exclusion. AI systems used exclusively for military, defense, or national security purposes fall outside the AI Act altogether. Commentators have noted that member states use national security exemptions to place surveillance practices beyond regulatory scrutiny.²⁰⁰

Separately, the GDPR prohibits the processing of biometric data for the purpose of uniquely identifying a person, and of data revealing political opinions, religious beliefs, or sexual orientation, subject to exceptions.²⁰¹ The Luxembourg Court has repeatedly held that the general and indiscriminate retention of traffic and location data is incompatible with the Charter, and that retention is permissible only where a member state faces a genuine threat, subject to review and proportionality limits.²⁰²

Data processing for criminal law enforcement purposes falls under the Law Enforcement Directive. This directive permits the processing of biometric and genetic data for identification only where strictly necessary and subject to certain safeguards. The Digital Omnibus on data, still under negotiation at the time of writing, would narrow the concept of personal data and reduce protections in other respects. The European Data Protection Board and the European Data Protection Supervisor have said the proposed changes would negatively affect the right to data protection.²⁰³

199 European Digital Rights, “The AI Act Isn’t Enough: Closing the Dangerous Loopholes That Enable Rights Violations,” <https://edri.org/our-work/the-ai-act-isnt-enough-closing-the-dangerous-loopholes-that-enable-rights-violations/>.

200 Ashwin Prabu and Marlena Wisniak, “When National Security Becomes a Shield for Evading AI Accountability,” *Tech Policy Press*, February 16, 2026, <https://www.techpolicy.press/when-national-security-becomes-a-shield-for-evading-ai-accountability/>. On the exclusion, see Artificial Intelligence Act art. 2(3).

201 Regulation (EU) 2016/679 (General Data Protection Regulation) art. 9(1); processing by competent authorities for criminal law enforcement purposes is excluded by art. 2(2)(d).

202 Joined Cases C-293/12 and C-594/12, *Digital Rights Ireland*, judgment of April 8, 2014; Joined Cases C-203/15 and C-698/15, *Tele2 Sverige and Watson*, judgment of December 21, 2016; Joined Cases C-511/18, C-512/18 and C-520/18, *La Quadrature du Net and Others*, judgment of October 6, 2020; Case C-623/17, *Privacy International*, judgment of October 6, 2020; Joined Cases C-793/19 and C-794/19, *SpaceNet and Telekom Deutschland*, judgment of September 20, 2022. On the narrower position regarding IP addresses, see Case C-470/21, *La Quadrature du Net II*, judgment of April 30, 2024.

203 EDPB-EDPS Joint Opinion 2/2026 on the proposal for a Digital Omnibus Regulation, adopted February 11, 2026.

Two developments are particularly worrying, because, while tackling real concerns, they could facilitate mass surveillance. The first is the temporary derogation from the confidentiality rules of the ePrivacy Directive, which permits providers of messaging and email services to scan communications voluntarily. That derogation lapsed in April 2026 and was restored in July 2026 by a regulation that applies until April 3, 2028.²⁰⁴ It covers not only known CSAM but also previously unseen material detected using AI, and the detection of grooming in conversations.

The second development is a permanent instrument being considered to replace this derogation, the proposed Child Sexual Abuse Regulation. If adopted, this regulation would require providers of a broad range of services to assess the risk that their products are used to distribute such material and to apply proportionate mitigation. In its original form, it would have permitted detection orders compelling providers to scan communications, including on end-to-end encrypted services. The Council removed mandatory detection from its negotiating mandate in November 2025, and negotiators have since agreed to protect encryption.²⁰⁵ The European Data Protection Board and the European Data Protection Supervisor have opposed the detection mandate since 2022, warning that it could lead to generalized and indiscriminate scanning and that client-side scanning breaks the security model of end-to-end encryption.²⁰⁶

204 Regulation (EU) 2026/1881 of the European Parliament and of the Council of 24 July 2026 on a temporary derogation from certain provisions of Directive 2002/58/EC as regards the use of technologies by providers of number-independent interpersonal communications services for the processing of personal and other data for the purpose of combating online child sexual abuse, arts. 3, 11, <https://eur-lex.europa.eu/eli/reg/2026/1881/oj/eng>. The previous derogation, Regulation (EU) 2021/1232, expired on 3 April 2026.

205 European Digital Rights, “The ‘Chat Control 1.0’ Saga,” August 4, 2026, <https://edri.org/our-work/the-chat-control-1-0-saga-big-tech-can-scan-our-private-messages-again-but-parliament-sent-a-strong-signal-against-mass-surveillance/>.

206 EDPB-EDPS Joint Opinion 04/2022 on the Proposal for a Regulation laying down rules to prevent and combat child sexual abuse, adopted July 28, 2022, esp. paras. 52 and 63.

11. MEASURES EMPOWERING FREEDOM OF EXPRESSION

11.1 Empowerment of Linguistic and Cultural Diversity

Governments can also use AI to empower expression and promote access to information. One way is to ensure that AI can be accessed and used in multiple languages. The EU has twenty-four official languages and more than sixty regional and minority languages.²⁰⁷

Major models are trained overwhelmingly on English-language data. The Commission's translation service examined Common Crawl, a widely used training corpus, and found Latvian at roughly 0.09 percent of the dataset, Irish at 0.07 percent and Maltese at 0.03 percent.²⁰⁸ The least-represented half of all EU official languages amounts to only 2.4 percent of the dataset.²⁰⁹ A speaker of one of those languages is likely to get worse results than an English speaker.

The EU has responded to this challenge in different ways. The Alliance for Language Technologies European Digital Infrastructure Consortium, with the Language Data Space, pools multilingual data across seventeen member states.²¹⁰ OpenEuroLLM and LLMs4EU are building open-source multilingual models intended to cover all official EU languages.²¹¹ Both began in 2025, and as of mid-2026 OpenEuroLLM had released datasets and reference models.²¹²

207 European Union, "Languages," https://european-union.europa.eu/principles-countries-history/languages_en.

208 European Commission, Directorate-General for Translation, "An EU Institutional LLM to Prevent Digital Language Extinction," https://translation.ec.europa.eu/language-data-and-ai-using-ai-break-down-language-barriers/eu-large-language-model-prevent-digital-language-extinction_en.

209 European Commission, Directorate-General for Translation, "An EU Institutional LLM."

210 European Commission, "Commission Welcomes New Initiative to Support European Cultural and Linguistic Diversity in Artificial Intelligence," press release, March 20, 2025, <https://digital-strategy.ec.europa.eu/en/news/commission-welcomes-new-initiative-support-european-cultural-and-linguistic-diversity-artificial>.

211 ALT-EDIC, "OpenEuroLLM," <https://www.alt-edic.eu/projects/openeurollm/> (project coordinated by Charles University, Czechia, and co-led by AMD Silo AI, Finland; ALT-EDIC contributes linguistic data creation and management, language technology evaluation and community building); ALT-EDIC, "LLMs4EU," <https://www.alt-edic.eu/projects/llms4eu/>.

212 OpenEuroLLM, "First Year Progress and Next Steps," March 6, 2026, <https://openeurollm.eu/blog/first-year-progress-and-next-steps>.

Under the Data Union Strategy, the Commission has announced a crowdsourcing initiative for language data in smaller European languages and the release of more than thirty million digitized works from Europe’s cultural institutions for AI training.²¹³ The Commission’s translation service has also published a human-translated benchmark that so far covers sixteen official languages.²¹⁴

In February 2026, the Commission also launched a Frontier AI Grand Challenge, which offers the winner strategic access to EuroHPC supercomputing capacity to train a frontier AI model. The Commission selected the EUROPA consortium, led by the Italian company Domyn, to build an open-source model of more than 400 billion parameters covering all twenty-four official EU languages.²¹⁵ A smaller precedent already exists in TildeOpen, a thirty-billion-parameter open model released in September 2025 by a Latvian company using two million GPU hours on the EuroHPC LUMI supercomputer.²¹⁶

11.2 Open Source

Open models are important for expression because they let users and organizations adapt a model to their own needs. They also grant independent researchers greater flexibility to examine how a model handles contested speech. In addition, they reduce the risk that a small number of firms hold decisive influence over the information most people access.²¹⁷

EU law gives open models a narrow advantage. Specifically, the AI Act exempts providers of GPAI models released under a free and open-source license from the requirements to keep technical documentation for the authorities and to supply information to downstream providers. The copyright policy and training-content summary obligations still apply. The exception also does “not apply to general-purpose AI models with systemic risk.”²¹⁸

213 European Commission, *Data Union Strategy: Unlocking Data for AI*, COM(2025) 835 final (November 19, 2025).

214 European Commission, Directorate-General for Translation, “An EU Institutional LLM to Prevent Digital Language Extinction.” The dataset is available at <https://huggingface.co/datasets/EC-DGT-AI/EU-MMLU> and via the European Language Data Space, <https://language-data-space.eu/catalogue/view-offer/e9cf10ca-44d8-4ecf-b2b0-83d998b0b9ca/>.

215 European Commission, “Commission Selects EUROPA Consortium as the Winner of the Frontier AI Grand Challenge, a Project to Build European Open-Source Frontier AI Model in All 24 EU Languages,” June 19, 2026, <https://digital-strategy.ec.europa.eu/en/news/commission-selects-europa-consortium-winner-frontier-ai-grand-challenge-project-build-european-open>.

216 Tilde, *TildeOpen LLM* (September 2025), trained on the EuroHPC LUMI supercomputer, <https://tilde.ai/tildeopen-llm/>.

217 Sayash Kapoor et al., “On the Societal Impact of Open Foundation Models,” Center for Research on Foundation Models, Stanford University, February 27, 2024, <https://arxiv.org/abs/2403.07918>.

218 Artificial Intelligence Act art. 53(2).

In June 2026, the Commission adopted an EU Open Source Strategy as part of a European Technological Sovereignty Package that also included a proposal for a Cloud and AI Development Act.²¹⁹ The Strategy aims to support the development, scaling, and long-term sustainability of open source across the public and private sectors.²²⁰ It asks public administrations to become anchor users and contributors through procurement guidance and open-source-friendly tendering. It commits to guidelines for the “fair assessment of open source bids.”²²¹

11.3 Media and AI Literacy

Media literacy—the ability to access, analyze, and engage with information effectively and safely²²²—is the clearest alternative to restricting expression. It addresses the concerns that drive disinformation policy by equipping people to handle disinformation themselves, rather than by deciding on their behalf what they may see. The high-level expert group convened by the Commission reported in 2018 that “[a]ny form of censorship either public or private should clearly be avoided,” and made media and information literacy one of its five pillars.²²³

The European Democracy Shield includes media and digital literacy under its third priority area. Its “Media Resilience Programme” aims to fund independent journalism and media literacy actions “reaching all age categories (youth, adults, seniors),” and the 2030 Roadmap on digital education is to cover “digital skills, AI literacy and critical thinking.”²²⁴

However, the Shield was announced in terms that give cause for concern. Presenting the Democracy Shield in 2024, the Commission President undertook to ensure that “malign information or propaganda (...) is swiftly removed and blocked.” Alexander Hohlfeld, an affiliate fellow at The Future of Free Speech, has observed that a commitment of that kind carries “the inherent risk of suppressing critical or divergent voices in the political

219 European Commission, “The EU Open Source Strategy,” adopted June 3, 2026, <https://digital-strategy.ec.europa.eu/en/policies/open-source-strategy>; European Commission, “Commission Proposes Tech Sovereignty Package to Strengthen Europe’s Digital Autonomy and Resilience,” June 3, 2026.

220 European Commission, “The EU Open Source Strategy.”

221 European Commission, “The EU Open Source Strategy.”

222 Tarja Laaninen, “Media Literacy: Fostering a Key Civic Skill in a Digital Information Environment,” EPRS Briefing PE 772.886 (May 2025): 1.

223 European Commission, Final Report of the High Level Expert Group.

224 European Commission and High Representative, Joint Communication, “European Democracy Shield,” November 12, 2025, 2, 18, 21–22.

discourse.”²²⁵ A democracy shield built on the premise that citizens cannot tell truth from falsehood invites an overly restrictive approach.

Separately, the Commission has published an AI Literacy Framework for primary and secondary education with the Organisation for Economic Co-operation and Development (OECD), and the OECD will assess Media and Artificial Intelligence Literacy in the 2029 round of the Programme for International Student Assessment.²²⁶ In addition, the AI Act requires providers and deployers to take measures to support the development of AI literacy among their staff and others operating their systems on their behalf, a duty the 2026 AI Omnibus softened.²²⁷

11.4 Accessibility and Inclusion

The AI Act requires providers of high-risk AI systems to ensure that those systems comply with the accessibility requirements of the Web Accessibility Directive and the European Accessibility Act.²²⁸ It also prohibits AI systems that exploit vulnerabilities due to age, disability, or a specific social or economic situation in a way that materially distorts behavior and causes or is reasonably likely to cause significant harm.²²⁹

225 Alexander Hohlfeld, “Europe Cannot Protect Democracy by Distrusting Its Citizens,” *The Bedrock Principle*, May 27, 2025, <https://www.bedrockprinciple.com/p/europe-cannot-protect-democracy-by>; see also The Future of Free Speech, feedback on the European Democracy Shield call for evidence, May 23, 2025.

226 OECD and European Union, *Empowering Learners for the Age of AI: An AI Literacy Framework for Primary and Secondary Education* (Paris: OECD Publishing, 2026), <https://doi.org/10.1787/65cd27d4-en>.

227 Artificial Intelligence Act art. 4, as amended by Regulation (EU) 2026/1744, applicable from July 27, 2026.

228 Artificial Intelligence Act art. 16(l).

229 Artificial Intelligence Act art. 5(1)(b).

12. CONCLUSION

Generative AI provides a unique opportunity to enhance freedom of expression and access to information. That opportunity comes, of course, with challenges and trade-offs. Legislation and policy should ensure that we take advantage of the promise of this technology while tackling the harms with well-scoped and proportionate rules.

A strong free expression framework is fundamental. The EU has robust written protections, which could constitute a solid basis if applied with a true commitment to our freedoms. The Charter and the Convention guarantee the right to receive and impart information, and the Strasbourg Court has long held that this protection extends to expression that offends, shocks, or disturbs.²³⁰ At the same time, governments have been too quick to justify restrictions when speech is controversial. Courts have often failed to check these excesses, particularly on hate speech. A technology that enables widespread circulation of expression—the internet—has also been viewed with concern, and that concern has been used to justify stricter controls. Neither the Strasbourg Court nor the Luxembourg Court has yet ruled on whether freedom of expression protects AI-generated content. Despite the strength of its democracies, the EU’s trajectory on contentious speech is concerning. As Europeans, we should all recommit to the pluralistic values that constitute a cornerstone of our democracies.

The EU has also adopted measures that deal with some of the issues arising for freedom of expression and that are a first step toward protecting our freedoms in this era. The AI Act prohibits the untargeted scraping of images to build facial recognition databases and contains a general prohibition on real-time remote biometric identification by law enforcement in publicly accessible spaces, subject to narrow exceptions. Both help protect public participation.

²³⁰ Charter of Fundamental Rights art. 11(1); *Handyside*, para. 49.

Still, concerns about AI have led to broad and vague rules. Obligations drafted in open terms that mix real issues, like cybersecurity, with vaguely defined worries like disinformation and broadly defined harms to society invite misguided enforcement, legal uncertainty, and self-censorship.

Ensuring that we realize the benefits of generative AI while tackling its harms is no easy task. Fortunately, the Charter and the Convention provide solid principles to determine how to handle it, if read with a true commitment to free expression and access to information. Restrictions should be prescribed by law with enough precision for people to know what is expected of them, and they should pursue a legitimate aim that we have all agreed on. Crucially, they should be necessary and proportionate. In other words, they should not go beyond what the objective requires, and they should be adopted only after less restrictive alternatives have been considered.²³¹ These rules, along with a commitment to the freedoms and pluralistic values that are the basis of our democracies, should serve us well in this new era.

²³¹ Charter of Fundamental Rights art. 52(1); Convention art. 10(2).



**THE
FUTURE
OF
FREE
SPEECH**

www.futurefreespeech.org