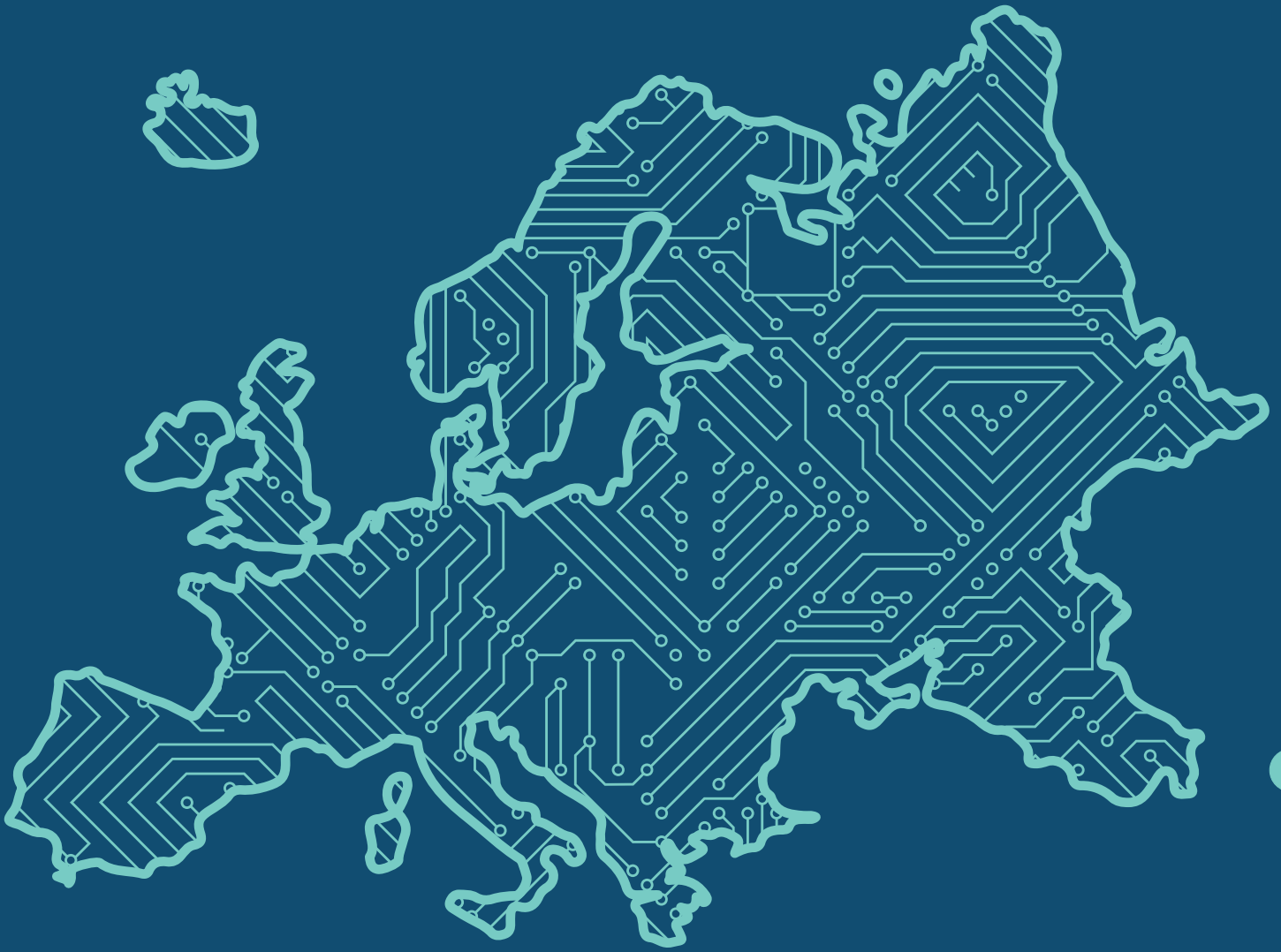




Artificial Intelligence and Freedom of Expression in the European Union

Jordi Calvet-Bademunt, Jacob Mchangama,
and Isabelle Anzabi*

* Jordi Calvet-Bademunt, Jacob Mchangama, and Isabelle Anzabi are a senior research fellow, executive director, and research associate, respectively, at The Future of Free Speech. We thank Natalie Alkiviadou, Joan Barata, and Alexander Hohlfeld for their valuable comments and suggestions. All remaining errors are our own.

Abstract

In this chapter we critically examine how generative artificial intelligence (AI) interacts with free speech standards in the European Union (EU). AI has the potential to be a powerful enabler of freedom of expression and access to information. While Article 10 of the European Convention on Human Rights (the Convention) and Article 11 of the EU Charter of Fundamental Rights establish strong protections in principle, the case law of the European Court of Human Rights (the Strasbourg Court) raises concerns in certain instances, especially regarding hate speech. The EU is not yet a party to the Convention, but this document exerts significant influence on EU law, a fact explicitly acknowledged in the EU Charter. The Strasbourg Court's recognition of states' discretion for broad national laws, combined with the poorly defined "systemic risk" provisions of both the AI Act and the Digital Services Act, could excessively limit freedom of expression and access to information in AI. Some elements of the EU framework are legitimate when narrowly applied, such as prohibiting AI-generated child sexual abuse material (CSAM) and recognizing exemptions for artistic and satirical works. However, the prevailing approach favors risk-aversion over robust speech protection, potentially creating structural pressures for over-blocking by providers and platforms. Although the EU is largely composed of mature democracies governed by the rule of law, the danger posed by the current AI governance approach to freedom of expression and access to information is a real concern. The EU faces a choice: Design AI governance that avoids excessive speech restrictions, or embed rules that could stifle legitimate voices and ideas.



Jordi Calvet-Bademunt

Jordi Calvet-Bademunt is a Senior Research Fellow at The Future of Free Speech. He is also a Visiting Legal Researcher at the Barcelona Supercomputing Center, where he advises on trustworthy AI. His work focuses on AI policy and digital governance, and he has written extensively and provided commentary in both specialist and mainstream media. Previously, Jordi spent about 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.



Jacob Mchangama

Jacob Mchangama is the Founder and Executive Director of The Future of Free Speech. He is a research professor at Vanderbilt University and a Senior Fellow at The Foundation for Individual Rights and Expression (FIRE). In 2018, he was a visiting scholar at Columbia's Global Freedom of Expression Center. He has commented extensively on free speech and human rights in outlets including the Washington Post, the Wall Street Journal, The Economist, Foreign Affairs and Foreign Policy. Jacob has published in academic and peer-reviewed journals, including Human Rights Quarterly, Policy Review, and Amnesty International's Strategic Studies. He is the producer and narrator of the podcast "Clear and Present" Danger: A History of Free Speech and the critically acclaimed book Free Speech: A History From Socrates to Social Media, published by Basic Books in 2022. He is the recipient of numerous awards for his work on free speech and human rights.



Isabelle Anzabi

Isabelle Anzabi is a research associate at The Future of Free Speech, where she analyzes the intersections between AI policy and freedom of expression. She is bringing her background in digital rights policy and global regulatory approaches to content moderation and AI governance. Previously, Isabelle was an AI & Human Rights Fellow with the European Center for Not-for-Profit Law, a Knowledge Fellow at the DiploFoundation, and a research group member at the Center for AI and Digital Policy. Isabelle received her B.A. in Political Science from Stanford University. She also studied digital governance at Oxford University and interned at institutions, such as the World Bank and CISA. On campus, Isabelle was affiliated with the Stanford Center for Racial Justice, the Stanford Legal Design Lab, the Stanford Cyber Policy Center, the Stanford Constitutional Law Center, the Stanford Technology Law Review, and the Public Service Leadership Program.

1. Introduction

Generative artificial intelligence (AI) has the capacity to revolutionize the way Europeans create, share, and access information. By making sophisticated tools for writing, translation, art, and research available to anyone, these systems can break down barriers of language, resources, and technical skill. They can empower vulnerable communities,¹ enable the preservation of cultural and linguistic diversity at scale,² and enrich democratic debate with diverse perspectives.³ In short, generative AI could be one of the most potent enablers of freedom of expression and access to information in the EU's history.

But Europe's current policy trajectory risks undermining these benefits before they can fully materialize. The EU's regulatory framework — anchored in the AI Act, the Digital Services Act (DSA), and a growing body of related policies — responds to some very concrete concerns (such as child sexual abuse material) and other much more diffuse and vague harms, like disinformation or hate speech. Some EU rules are drafted in sweeping, ambiguous terms that could lead to overreach.⁴ Provisions on “systemic risks” and content moderation obligations are so broadly framed that they could justify the restriction of lawful, even vital, speech — whether political criticism, satire, or reporting that challenges prevailing narratives.

This situation is compounded by the ambivalent protection of the right to freedom of expression displayed by the European Court of Human Rights (the Strasbourg Court)⁵ — which applies the European Convention on Human Rights (the Convention). While the EU is not yet a party to the Convention,⁶ the Convention nonetheless exerts significant influence on EU law, as explained in the following section. As a result, the Convention remains a crucial reference point for the protection of fundamental rights within the EU.

The Strasbourg Court has traditionally recognized that freedom of expression protects information and ideas that offend, shock, or disturb;⁷ however, more recent case law and the EU's recent legislative trend in the AI era risk eroding this principle. By placing broad risk-based provisions and encouraging preemptive moderation, lawmakers chance pushing providers and platforms into over-moderation as a defensive measure. This “better safe than sorry” approach may protect against some harms, but it could also chill the creativity, pluralism, and political discourse that are essential to a democratic society.

A comparable pattern can be seen with social media. The former United Nations Special Rapporteur on Freedom of Opinion and Expression cautioned against imposing obligations on companies to restrict content

1 Stefano Regondi, Giordana Donvito, Emanuele Frontoni, Milutin Kostovic, Fabio Minazzi, Sébastien Bratières, Massimiliano Filosto, and Raffaele Pugliese, “Artificial Intelligence Empowered Voice Generation for Amyotrophic Lateral Sclerosis Patients,” *Scientific Reports* 15, no. 1361 (2025): 1–12, <https://doi.org/10.1038/s41598-024-84728-y>.

2 Adeyinka Tella, Esther Oluwayemi Jatto, and Yusuf Ayodeji Ajani, “Preserving Indigenous Knowledge: Leveraging Digital Technology and Artificial Intelligence,” *IFLA Journal* 51, no. 2 (2025), <https://doi.org/10.1177/03400352251342505>.

3 Chris Stokel-Walker, “A Partnership Between Jigsaw and This Kentucky City Could Be the Future of Civic Engagement,” *Fast Company*, February 14, 2025, <https://www.fastcompany.com/91278879/jigsaw-bowling-green-kentucky-civic-engagement>.

4 Jordi Calvet-Bademunt, “Safeguarding Freedom of Expression in the AI Era,” *Tech Policy Press*, November 4, 2024, <https://www.techpolicy.press/safeguarding-freedom-of-expression-in-the-ai-era>.

5 The Strasbourg Court is different from the Court of Justice of the EU. The Court of Justice of the EU interprets EU law and settles legal disputes between national governments and EU institutions. The Strasbourg Court was established by the ECHR, outside the EU framework, and rules on issues related to this Convention.

6 Nicoletta Ionta, “Commission Seeks EU Court Nod for ECHR Bid After Years of Gridlock,” *Euractiv*, July 25, 2025, <https://www.euractiv.com/section/politics/news/commission-seeks-eu-court-nod-for-echr-bid-after-years-of-gridlock>.

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

on the basis of vague or complex legal standards when such measures lack prior judicial oversight and carry the threat of severe penalties.⁸ This is arguably what some of the rules governing generative AI do. The Special Rapporteur highlighted that these rules could endanger freedom of expression by pressuring companies to err on the side of over-removal, leading them to take down lawful content to shield themselves from liability.⁹ The stakes are particularly high because generative AI is quickly becoming a core part of how information is produced and consumed. If lawful AI-generated expression is routinely suppressed or if public debate is filtered through vague risk-based controls, the EU could see a narrowing of its informational and cultural space at precisely the moment it has the tools to expand it. Policymakers must therefore resist the temptation to regulate away uncertainty by granting themselves or companies open-ended discretion to decide what forms of speech are permissible.

If the EU truly intends to be a global leader in both technology governance and human rights, it must approach the governance of generative AI with rigor and restraint. That means ensuring any restriction is prescribed by law, legitimate, and necessary. It should also be evidence-based — not a product of political expediency or generalized fear of the unknown. Anything less risks turning a transformative engine of expression into a tightly managed channel of approved speech.

⁸ Human Rights Council: Special Rapporteur on Freedom of Opinion and Expression, “Report on Content Regulation,” A/HRC/38/35, para. III.A.15 (April 6, 2018), <https://www.ohchr.org/en/calls-for-input/report-content-regulation>.

⁹ Special Rapporteur, “Rep. on Content Regulation,” A/HRC/38/35, para. III.A.17.

2. Substantive Analyses

2.1. General Standards of Freedom of Expression

European fundamental rights protect freedom of expression and access to information while allowing for restrictions in specific cases. The EU Charter of Fundamental Rights, the Convention, and national laws constitute the legal framework protecting and promoting freedom of expression in the EU.

The EU Charter provisions are addressed to EU institutions and bodies and national authorities only when they are implementing EU law.¹⁰ Though the Convention is not an EU document, it applies to member states because they have all ratified it. In addition, the EU Charter of Fundamental Rights, which is an EU document, explicitly provides that, insofar as it contains rights corresponding to those guaranteed by the Convention, the meaning and scope of those rights must align with those laid down in the Convention or be more protective than Convention rights.¹¹ Furthermore, the Treaty of the EU recognizes the Convention fundamental rights as general principles of EU law.¹² As a result, the Convention exerts significant influence on EU law.

Article 10 of the Convention guarantees the right to freedom of expression, including the freedom to hold opinions and to receive and impart information regardless of frontiers. The Strasbourg Court has affirmed that this protection extends even to information or ideas that “offend, shock or disturb.”¹³ Similarly, Article 11 of the EU Charter of Fundamental Rights protects freedom of expression and information.

Freedom of expression is not absolute. The Convention permits certain restrictions “prescribed by law” and “necessary in a democratic society”¹⁴ for legitimate aims, such as protecting national security,¹⁵ public safety,¹⁶ health or morals,¹⁷ or the rights and reputations of others.¹⁸ In a similar vein, the Charter points out that any limitations to these rights must be “provided for by law and respect the essence of those rights and freedoms.”¹⁹ The Charter also subjects limitations to the principle of proportionality; they must be “necessary and genuinely meet objectives of general interest ... or the need to protect the rights and freedoms of others.”²⁰

¹⁰ Charter of Fundamental Rights of the European Union, Official Journal of the European Union [O.J.] (C 326), art. 52 (2012): 391-407, <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:12012P/TXT>.

¹¹ Charter of Fundamental Rights of the E.U., art. 52.

¹² Treaty on European Union, O.J. (C 202), art. 6, para. 3 (2016): 19, <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=OJ:C:2016.202.FULL>.

¹³ *Handyside v. U.K.* (1976).

¹⁴ 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>.

¹⁵ 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>.

¹⁶ See, *inter alia*, *Information Note on the Court's Case-Law No. 112: 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>.

¹⁷ 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>.

¹⁸ See, *inter alia*, *Information Note on the Court's Case-Law No. 121: Féret v. Belgium*, App. No. 15615/07, Eur. Ct. H.R. (July 16, 2009), <https://hudoc.echr.coe.int/eng?i=002-1407>; *Information Note on the Court's Case-Law No. 64: Seurot v. France (dec)*, App. No. 57383/00, Eur. Ct. H.R. (May 18, 2004), <https://hudoc.echr.coe.int/eng?i=002-4404>; *Vejdeland 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>.

¹⁹ Charter of Fundamental Rights of the E.U., art. 52.

²⁰ Charter of Fundamental Rights of the E.U., art. 52.

The Strasbourg Court and the Court of Justice of the European Union (the Luxembourg Court) have not yet ruled on AI-generated content. Still, there are reasons to consider that the right to freedom of expression should extend to users creating content with AI, as has been suggested by the Council of Europe’s committee of experts on the implications of generative AI for freedom of expression. The expert group’s preliminary analysis suggests that factors such as whether the expression is generated under an individual’s agency and the extent to which the AI output reflects the user’s intent are relevant considerations.²¹

In our view, the fundamental right to access information should also protect citizens who wish to access content autonomously generated by AI. Some scholars have pointed out that “the listener’s standpoint may more accurately capture the sense of extending constitutional coverage to generative AI output.”²² This will become increasingly important in an environment where AI contributes more and more to knowledge creation and dissemination, for instance, through AI-powered search engines like Google’s AI Mode and ChatGPT search feature. Users should not be disadvantaged simply because the source of the content is an AI model rather than a human author. The EU should also focus on cultivating a thriving environment for open models, reducing the risk that a few companies have an outsized influence on the ecosystem of ideas and information. Open models empower greater customizability, mitigate the harm of monoculture, and reduce market concentration.²³

Any state measure that broadly bans, censors, or limits access to AI outputs should be considered under the usual legality, legitimacy, and necessity tests. Interferences must be prescribed by law, pursue a legitimate aim, and be narrowly tailored to avoid undue chilling of speech. Overbroad bans on controversial AI-generated statements, for example, could raise serious concerns for free expression, potentially suppressing debate or artistic creativity.

The Council of Europe has convened a Committee of Experts on Generative Artificial Intelligence for Freedom of Expression (MSI-AI), whose role is specifically to examine the impacts of generative AI on freedom of expression and to draft guidance by the end of 2025.²⁴ The Council of Europe’s Framework Convention on Artificial Intelligence and Human Rights, Democracy, and the Rule of Law was opened for signature in September 2024.²⁵ At the time of writing, it has 16 signatories, including the European Union and the United States.²⁶

2.2. AI-Specific Legislation and Policies

The European Union has taken an early and leading role in regulating AI, with the EU’s AI Act at the forefront of global efforts. The AI Act adopts a risk-based approach: It bans a few extreme use cases deemed to pose unacceptable risk, heavily regulates “high-risk” applications, and imposes transparency requirements on more general systems with limited risk. This includes informing users that they are interacting with AI in chatbots and obligations to watermark and label deepfakes.

21 Council of Europe: Committee of Experts on the Implications of Generative Artificial Intelligence for Freedom of Expression (MSI-AI), “Draft Guidance Note on the Implications of Generative Artificial Intelligence for Freedom of Expression,” MSI-AI(2025)10 (June 3, 2025): 8–9, <https://rm.coe.int/msi-ai-2025-10-draft-guidance-note-on-the-implications-of-generative-a/1680b68c48>.

22 Marco Bassini, “Speech Without a Speaker: Constitutional Coverage for Generative AI Output,” *European Constitutional Law Review* (July 31, 2025): 23, <https://doi.org/10.1017/S1574019625100771>.

23 Sayash Kapoor et al., “On the Societal Impact of Open Foundation Models: Analyzing the Benefits and Risks of Foundation Models with Widely Available Weights,” Center for Research on Foundation Models, Stanford University, February 27, 2024, <https://crfm.stanford.edu/open-fms>.

24 Council of Europe, “First Meeting of the Council of Europe Committee of Experts on the Implications of Generative Artificial Intelligence for Freedom of Expression (MSI-AI) Held in Strasbourg April 23–24, 2024,” accessed August 28, 2025, <https://www.coe.int/en/web/freedom-expression/-/first-meeting-of-council-of-europe-committee-of-experts-on-the-implications-of-generative-artificial-intelligence-on-freedom-of-expression-msi-ai-held-in-strasbourg>.

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

26 Council of Europe, CETS No. 225 (2024).

Generative AI has been a focus of lawmakers' attention, especially after the rapid emergence of systems like ChatGPT. In its final form, the AI Act introduced several provisions directly relevant to generative AI. These rules apply to the so-called general-purpose AI (GPAI) models — AI models, such as OpenAI's GPT-5, trained on large amounts of data that can perform multiple tasks effectively.

The requirements for GPAI models include maintaining comprehensive technical documentation, ensuring compliance with EU copyright laws, and publishing summaries of the training data used. For GPAI models identified as posing systemic risks — such as those trained with exceptionally large computational resources — additional requirements apply: conducting standardized model evaluations, assessing and mitigating systemic risks, reporting serious incidents to authorities, and implementing robust cybersecurity measures. Rules governing systemic risk have raised concerns for freedom of expression due to their vagueness.²⁷ Over 20 entities (including leading AI providers such as Anthropic, OpenAI, Google, and Mistral AI) have signed a Code of Conduct for GPAI models that further develops the obligations introduced by the AI Act.

The act provides a limited exemption for GPAI models released under free and open-source licenses that allow access, usage, modification, and distribution. To qualify, these models must make their parameters, including weights, architecture, and usage information, publicly available. Even then, providers must still produce a summary of the training data and implement policies to ensure compliance with EU copyright laws. Notably, this exemption does not apply to GPAI models deemed to present systemic risks, such as those with high computational capabilities, which are subject to the same obligations as closed models presenting systemic risks.

Liability for AI-generated content is another emerging issue. The European Commission proposed an AI Liability Directive to complement the AI Act. This directive would have allowed victims of damage caused by AI (including reputational harm) to more easily sue for compensation. But the European Commission ultimately withdrew its proposal.²⁸ The AI Liability Directive could have raised concerns about freedom of expression, as it might have incentivized companies to preemptively censor their models to avoid liability. Still, the EU updated its product liability rules, effective December 2026, to make clear that software and AI can trigger liability when they malfunction and cause harm, just like physical products.²⁹ The new rules classify software and AI as “products,” holding developers and providers strictly liable for defects, even those arising post-sale through updates or machine learning. The Product Liability Directive broadens the scope of compensable damages to include medically recognized psychological harm and data loss, eases the burden of proof for claimants, and extends liability timelines up to 25 years for latent injuries. This framework complements the AI Act by ensuring that individuals harmed by AI technologies have clear legal avenues for compensation.³⁰ It is crucial that liability rules are interpreted in a manner that does not chill access to legitimate information by prompting companies to withhold content out of liability concerns.

The EU has also signed the Council of Europe's Framework Convention on AI, Human Rights, Democracy and the Rule of Law, although it has not yet ratified it. This convention is the first binding international treaty on AI. It applies to activities within the life cycle of AI systems undertaken by public authorities or private

27 Calvet-Bademunt, “Safeguarding Freedom of Expression in the AI Era.”

28 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>.

29 Lena Niehoff, David Hilger, Megan Howarth, and Katie Chandler, “New Product Liability Directive 2024/2853: New Product Liability Risks for Products in the EU,” Taylor Wessing Insights, January 6, 2025, <https://www.taylorwessing.com/en/insights-and-events/insights/2025/01/di-new-product-liability-directive>.

30 Stefano De Luca, “EU Legislation in Progress: Revised Product Liability Directive,” EPRS Briefing PE 739.341 (February 19, 2025), https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/739341/EPRS_BRI%282023%29739341_EN.pdf.

actors acting on their behalf. Parties to the convention must also address risks and impacts arising from private actors, but they have greater flexibility in how to do so. The convention outlines seven core principles, including human dignity and individual autonomy, transparency, accountability, and privacy. It also establishes obligations to protect human rights, safeguard the integrity of democratic processes, and uphold respect for the rule of law.

In the EU, generative AI may also be indirectly regulated by the DSA, which imposes broad regulation on online platforms and search engines. The DSA requires very large online platforms (VLOPs) and very large online search engines (VLOSEs) to assess and mitigate systemic risks. Although not identical, these risks are similar to those outlined in the AI Act. EU officials have hinted that the DSA will govern AI features on online platforms and AI-powered search engines.³¹ As with the AI Act, these systemic risk obligations can raise concerns about freedom of expression due to their vagueness and breadth.³²

Under the DSA, online platforms benefit from a liability exemption for third-party content they host, provided they act expeditiously to remove illegal material upon notification. There is an ongoing debate about whether this exemption protects providers of AI services. It is possible that certain AI services,³³ such as generative AI integrated into search engines, would qualify for the exemption.³⁴ However, the exemption would not apply when AI generates content more autonomously. In such cases, the AI provider or deployer would face direct liability. While this issue remains the subject of ongoing debate, imposing an overly strict liability regime for generative AI could prompt companies to restrict content excessively — for example, through the use of aggressive filters. This appears to have happened when specific names, such as “Brian Hood,” caused ChatGPT to end conversations abruptly a few months ago. Brian Hood is the name of an Australian mayor who threatened to sue OpenAI for defamation after ChatGPT generated some false statements about him.³⁵ Experts suggest that OpenAI’s solution was to hardcode ChatGPT to refuse generating content about specific personal names.³⁶

2.3. Defamation

Generative AI systems can and sometimes do produce false statements about individuals — including content that may be defamatory. This raises a question: If an AI chatbot states, for instance, that someone committed a crime (when in reality they did not), who is legally responsible for the harm caused? These new scenarios are testing European laws on defamation. The challenge is to fit AI-generated content into the traditional framework that establishes liability between primary publishers, intermediaries, and re-publishers of defamatory material.

There is no single EU-wide defamation law; instead, each member state applies its own legal framework, shaped by national history and its interpretation of how to balance reputation with freedom of expression under the European Convention on Human Rights. The Convention, as interpreted by the Strasbourg Court,

31 See European Parliament, “Answer Given by Executive Vice-President Virkkunen on Behalf of the European Commission,” Parliamentary Question E-000394/2025(ASW) (April 24, 2025), https://www.europarl.europa.eu/doceo/document/E-10-2025-000394-ASW_EN.html; Luca Bertuzzi (@BertuzLuca), “EXCL: The EU Commission is considering designating ChatGPT as a systemic platform under DSA due to its web-searching functionality. OpenAI recently reported a jump from 11.2 m to 41.3 m monthly active users in the EU, nearing the 45 m threshold,” X, April 30, 2025, <https://x.com/BertuzLuca/status/1917609828987326743>.

32 Calvet-Bademunt, “Safeguarding Freedom of Expression in the AI Era.”

33 Laureline Lemoine and Mathias Vermeulen, “Assessing the Extent to Which Generative Artificial Intelligence (AI) Falls Within the Scope of the EU’s Digital Services Act: An Initial Analysis,” SSRN, October 9, 2023, <https://ssrn.com/abstract=4702422>.

34 Laura Pliauskaitė and Uzma Chaudhry, “AI and Digital Governance: Exploring Platform Liability Laws in the EU,” IAPP, September 25, 2024, <https://iapp.org/news/a/ai-and-digital-governance-exploring-platform-liability-laws-in-the-eu>.

35 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>.

36 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>.

generally affords weaker protection to freedom of expression than the US First Amendment, particularly in cases involving political speech and defamation of public officials and figures.³⁷ Many countries still criminalize defamation, including provisions for imprisonment.³⁸ The Council of Europe has cautioned that overly protective defamation laws can have a chilling effect on public debate and has urged member states to decriminalize defamation.³⁹ Nevertheless, in many jurisdictions within the EU, such laws continue to raise serious concerns for free expression.

Some scholars suggest that “the coming of the Internet has determined an increase of the consideration given to restrictions to free speech” by the Strasbourg Court, reflecting a perceived amplification of online threats.⁴⁰ The Strasbourg Court has explicitly stated that the risk of harm from internet content is “certainly higher” than from the press.⁴¹ Although the Strasbourg Court still polices disproportionate measures,⁴² its decisions, taken as a whole, show a growing readiness to justify restrictions on speech. This is illustrated by cases in which the Strasbourg Court found no violation of Article 10 of the Convention, such as when Estonia held an online news portal responsible for defamatory remarks left by users in its comment section,⁴³ or when France convicted a politician for failing to promptly remove unlawful third-party comments from the public wall of his Facebook account.⁴⁴

While there is a lack of court decisions on whether people can seek injunctions or damages for false or defamatory statements made by AI,⁴⁵ the overly restrictive approach applied to the internet may also be applied to AI. This would be bad news for freedom of expression and for AI’s promise as an unprecedented tool to access information, as it would incentivize companies to err on the side of caution and over-censor. As explained in the previous section, this is not a purely hypothetical concern. Instances exist of AI providers applying broad limits to the content they display, presumably in response to potential liability for defamation.⁴⁶

At the time of writing, no clear national or EU laws specifically address who is liable for AI-generated content.⁴⁷ Still, according to some scholars, AI providers may be considered liable for untrue factual claims or defamatory speech about a person, at least in some member states.⁴⁸ Notably, while case law is not harmonized across the EU,⁴⁹ 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.⁵⁰ A similar reasoning could be applied to generative AI, should it generate defamatory content.

Importantly, when AI systems like ChatGPT generate new content, rather than just retrieving content as a search engine,⁵¹ they are highly unlikely to benefit from the “safe harbors” in the DSA that exempt some

37 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 (Edward Elgar Publishing, 2014): 541, and in Bocconi Legal Studies Research Paper No. 2706112 (December 22, 2015), <https://ssrn.com/abstract=2706112>.

38 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.

39 Council of Europe, “Freedom of Expression: Defamation,” accessed August 28, 2025, <https://www.coe.int/en/web/freedom-expression/defamation>.

40 Pollicino and Bassini, “Free Speech, Defamation and the Limits to Freedom of Expression in the EU,” 530.

41 Editorial Board of *Pravoye Delo* and *Shtekel v. Ukraine*, App. No. 33014/05, Eur. Ct. H.R. (May 5, 2011, final August 5, 2011), HUDOC no. 001-104685; discussed in Pollicino and Bassini, “Free Speech, Defamation and the Limits to Freedom of Expression in the EU.”

42 Ahmet Yildirim v. Turkey, App. No. 3111/10, Eur. Ct. H.R. (December 18, 2012, final March 18, 2013), HUDOC no. 001-115705, discussed in Pollicino and Bassini, “Free Speech, Defamation and the Limits to Freedom of Expression in the EU.”

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

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

45 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/oxfordhb/9780198940272.013.0022>.

46 Masnick, “Curious Case of ChatGPT’s Banned Names.”

47 Kastl-Riemann, “Regulation of Generative AI Speech.”

48 Kastl-Riemann, “Regulation of Generative AI Speech.”

49 Stavroula Karapapa and Maurizio Borghi, “Search Engine Liability for Autocomplete Suggestions: Personality, Privacy and the Power of the Algorithm,” *International Journal of Law and Information Technology* 23, no. 3 (Autumn 2015): 261–89, <https://doi.org/10.1093/ijlit/eaav009>.

50 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>.

51 Pliuškaitė and Chaudhry, “AI and Digital Governance.”

services, such as online platforms, from liability.⁵² These protections do not extend to AI-generated content because the output is not coming from a third-party user but rather is created by the AI itself.⁵³ 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.⁵⁴

There is also the question of the user's liability when re-publishing AI statements. If someone uses a chatbot and then posts its answer publicly, that user becomes a publisher of the information. In many member states, such as France⁵⁵ and Germany,⁵⁶ repeating a defamatory allegation is treated as equivalent to making it in the first place. That said, users may rely on existing defenses — such as proving that there was a factual basis for their allegations or that they were acting in good faith.⁵⁷

European law offers a potential alternative route in relation to false information: data protection (privacy) law. Suppose the false statement contains personal data, which is likely to be the case in a defamation case. In such cases, the General Data Protection Regulation (GDPR) provides individuals with the right to correct or delete inaccurate personal data. The GDPR route, however, has limitations. Notably, it entitles the affected person to rectification but not compensation.⁵⁸ Some organizations, such as the privacy advocacy group Noyb, have filed complaints against OpenAI, assisting citizens who argue that ChatGPT processes false information about them and, therefore, violates the GDPR.⁵⁹

2.4. Explicit Content

Generative AI has introduced serious new challenges for child sexual abuse material (CSAM) and nonconsensual intimate imagery (NCII). AI systems can create realistic images or videos depicting child sexual abuse without involving an actual child. Likewise, generative AI can produce “deepfake” pornography, superimposing a person's (adult) face onto explicit sexual material without consent. These developments challenge law enforcement and raise complex questions about how freedom of expression doctrines apply to such content.

52 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>.

53 Arcila, “Is It a Platform?,” 460.

54 Kastl-Riemann, “Regulation of Generative AI Speech.”

55 Dalloz, “Fiches d'orientation: Diffamation,” September 2022, <https://www.dalloz.fr/documentation/Document?id=DZ/OASIS/000328>; Anticor, “La diffamation sur internet,” April 10, 2025, <https://www.anticor.org/outils-citoyens/comment-eviter-la-diffamation-sur-internet>.

56 Strafgesetzbuch (Penal Code), §§ 186–87 (2021), https://www.gesetze-im-internet.de/englisch_stgb/englisch_stgb.html (Ger.).

57 Council of Europe: Eur. Ct. H.R., “Guide on Article 10 of the European Convention on Human Rights: Freedom of Expression,” para. 232–37 (updated August 31, 2022), <https://rm.coe.int/guide-on-article-10-freedom-of-expression-eng/native/1680ad61d6>.

58 Kastl-Riemann, “Regulation of Generative AI Speech.”

59 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>.

2.4.1. CSAM

Europol — the EU’s agency for Law Enforcement Cooperation — has recently reported an increase in AI-facilitated CSAM cases and has been involved in arrests worldwide.⁶⁰ Offenders use generative AI in two main ways: (1) to create “deepfakes” by inserting real children’s faces from innocuous photos into sexual images, and (2) to generate entirely new abuse images of fictional children who do not exist.⁶¹ This reflects a global issue. The National Center for Missing and Exploited Children received 485,000 reports of AI-related CSAM in the first half of 2025, compared with 67,000 for all of 2024.⁶²

Under EU law, CSAM is strictly outlawed under Directive 2011/93/EU, which combats the sexual abuse and sexual exploitation of children and child pornography. The directive includes “realistic images of children” within its definition of “child pornography,” using forward-looking language that potentially covers virtual or computer-generated material and makes it punishable. However, it also allows member states to limit criminalization to depictions of real and existing children, creating discrepancies between European countries.⁶³

The Strasbourg Court has stated that CSAM is not protected under the right to freedom of expression. In *Karttunen v. Finland* (2011), the Strasbourg Court reviewed the conviction of an artist who included child pornography in an exhibition. While acknowledging an interference with her freedom of expression under Article 10 of the Convention, the Strasbourg Court found the interference to be justified and proportionate. It emphasized the necessity of protecting children and their privacy rights and upholding public morals, ruling the case inadmissible as manifestly ill-founded.⁶⁴

Europe may be moving toward even more vigorous enforcement through new legislation. In 2022, the European Commission proposed a Child Sexual Abuse Regulation.⁶⁵ Under the proposed regulation, a range of services — including hosting providers and publicly available interpersonal communications services (e.g., messaging and email) — could be subject to detection orders where a competent authority finds a significant risk. These orders could require providers to deploy technologies to detect and report known CSAM, previously unseen CSAM, and grooming, while removal (for hosting services) and blocking (for internet access providers) could be ordered separately. Compliance with detection orders could entail messaging services, such as WhatsApp and Signal, monitoring their services for CSAM. The Parliament highlighted the need to exclude end-to-end encryption from the scope of detection orders to ensure that these communications remain secure and confidential.⁶⁶ The regulation is currently under discussion. Ten countries, including Germany and the Netherlands, blocked an earlier version in 2024, primarily due to concerns that it was overly pro-surveillance.⁶⁷ Given its broad scope, the European Data Protection Board and the European Data Protection Supervisor have warned that the proposed regulation could lead to “generalized and indiscriminate scanning”

60 See Mike Corder, “AI Is Turbocharging Organized Crime, EU Police Agency Warns,” Associated Press, March 18, 2025, <https://apnews.com/article/europe-crime-europol-ai-security-cyber-attack-846847536f6feb2bbb423943fd96e1f1>; “25 Arrested in Global Hit Against AI-Generated Child Sexual Abuse Material,” Europol, February 28, 2025, <https://www.europol.europa.eu/media-press/newsroom/news/25-arrested-in-global-hit-against-ai-generated-child-sexual-abuse-material>.

61 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): 67, <https://doi.org/10.3390/laws13060067>.

62 Cecilia Kang, “A.I.-Generated Images of Child Sexual Abuse Are Flooding the Internet,” *New York Times*, July 10, 2025, updated July 18, 2025, <https://www.nytimes.com/2025/07/10/technology/ai-csam-child-sexual-abuse.html>.

63 Parti and Szabó, “Legal Challenges of Realistic and AI-Driven Child Sexual Abuse Material,” 11.

64 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>.

65 Proposal for a Regulation of the European Parliament and of the Council Laying Down Rules to Prevent and Combat Child Sexual Abuse, COM (2022) 209 final (May 11, 2022), <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022PC0209>.

66 European Parliament, “Child Sexual Abuse Online: Effective Measures, No Mass Surveillance,” press release, November 14, 2023, <https://www.europarl.europa.eu/news/en/press-room/20231110IPRI0118/child-sexual-abuse-online-effective-measures-no-mass-surveillance>.

67 Sam Clark, “Denmark Faces Uphill Battle on EU Child Sexual Abuse Bill,” *Politico*, June 3, 2025, <https://www.politico.eu/article/denmark-online-protections-minors-child-sexual-abuse-whatsapp-signal>.

of electronic communications, posing serious threats to privacy, data protection, and freedom of expression.⁶⁸ As of the time of writing, Denmark is proposing to reinsert controversial mandatory detection orders, forcing the scanning of encrypted messages.⁶⁹

There have also been significant concerns regarding the use of CSAM to train AI models. In 2023, Stanford University researchers “found hundreds of known images of [CSAM] in an open dataset used to train popular AI image generation models, such as Stable Diffusion.”⁷⁰ The EU’s AI Act, as currently implemented, requires model providers to describe the measures adopted to avoid or remove illegal content from training data.⁷¹

2.4.2. NCII

NCII can falsely depict individuals engaging in sexual acts they never performed. Such technology has been widely exploited to humiliate and harass women. Research shows that over 90% of deepfake content online is pornographic, with women — especially those in cultural, media, and political spheres — constituting the overwhelming majority of victims.⁷²

Until recently, many European jurisdictions did not explicitly criminalize deepfakes, leaving victims to rely on related offenses, such as violations of privacy laws, to seek justice.⁷³ The GDPR grants individuals rights over their data and allows sanctions for unauthorized publication. However, its effectiveness in addressing gender-based abuse has been limited. Victims often face cumbersome procedures, jurisdictional challenges with non-EU platforms, and persistent non-cooperation from certain services, including some messaging apps and pornographic websites.⁷⁴

In May 2024, the EU adopted Directive 2024/1385 to combat violence against women and domestic violence. This law sets minimum standards for criminalizing cyberviolence against women, including the nonconsensual creation and sharing of intimate or manipulated material.⁷⁵ The directive explicitly states that this should include “the fabrication of ‘deepfakes,’ where the material appreciably resembles an existing person ... and would falsely appear to other persons to be authentic.”⁷⁶

The new EU directive reflects a consensus that nonconsensual intimate deepfakes are a serious violation of sexual privacy and dignity. Thus, criminalizing NCII is seen as a justified restriction on expression. Of course, laws must be carefully worded to avoid capturing consensual uses (e.g., consensual role-playing with AI) or obvious satire. The focus is on intimate images and on lack of consent and harmful intent or effect.

68 European Data Protection Board and European Data Protection Supervisor, “Proposal to Combat Child Sexual Abuse Online Presents Serious Risks for Fundamental Rights,” joint press release, July 29, 2022, https://www.edps.europa.eu/press-publications/press-news/press-releases/2022/combat-child-sexual-abuse-online-presents-serious-risks-fundamental-rights_en.

69 Claudie Moreau, “Danish Presidency Brings Back Controversial Detection Orders in Child Abuse Law,” *Euractiv*, July 2, 2025, <https://www.euractiv.com/section/tech/news/danish-presidency-brings-back-controversial-detection-orders-in-child-abuse-law>.

70 David Thiel, “Investigation Finds AI Image Generation Models Trained on Child Abuse,” *Stanford Internet Observatory News*, December 20, 2023, <https://cyber.fsi.stanford.edu/news/investigation-finds-ai-image-generation-models-trained-child-abuse>.

71 Paul, Weiss Rifkind, Wharton & Garrison LLP, “EU Commission Publishes Guidelines on General-Purpose AI Obligations as Well as Training Data Disclosure Template: Further Clarity as the Countdown to Enforcement Begins,” client memorandum, July 30, 2025, <https://www.paulweiss.com/insights/client-memos/eu-commission-publishes-guidelines-on-general-purpose-ai-obligations-as-well-as-training-data-disclosure-template-further-clarity-as-the-countdown-to-enforcement-begins>.

72 Ionel Zamfir and Colin Murphy, “Cyberviolence Against Women in the EU,” *EPRIS Briefing PE 767146* (December 2024), https://www.europarl.europa.eu/RegData/etudes/BRIE/2024/767146/EPRS_BRI%282024%29767146_EN.pdf.

73 Karolina Mania, “Legal Protection of Revenge and Deepfake Porn Victims in the European Union: Findings from a Comparative Legal Study,” *Trauma, Violence, & Abuse* 25, no. 1 (2022), <https://doi.org/10.1177/15248380221143772>; Carlotta Rigotti and Clare McGlynn, “Towards an EU Criminal Law on Violence Against Women: The Ambitions and Limitations of the Commission’s Proposal to Criminalise Image-Based Sexual Abuse,” *New Journal of European Criminal Law* 13, no. 4 (2022): 452–77, <https://doi.org/10.1177/20322844221140713>.

74 Zamfir and Murphy, “Cyberviolence Against Women in the EU,” 6.

75 Zamfir and Murphy, “Cyberviolence Against Women in the EU,” 6.

76 Directive 2024/1385 on Combating Violence Against Women and Domestic Violence, 2024 O.J. (L 1385) (May 14, 2024), <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024L1385>.

The DSA also offers an avenue to tackle NCII. The DSA requires providers of VLOPs and online search engines to assess and mitigate the “systemic risks.” Notably, the DSA explicitly states that platforms primarily used for the dissemination to the public of pornographic content, including the nonconsensual sharing of intimate or manipulated material, should rapidly process notices and remove such content.⁷⁷ In May 2025, the European Commission opened formal proceedings against Pornhub, Stripchat, XNXX, and XVideos, which so far focused on the protection of minors and not on NCII.⁷⁸

2.5. Hate Speech

Hate speech is another category of expression subject to legal restrictions under European and international standards, and generative AI systems can potentially produce hateful content or be used to amplify it. The EU has an established framework dealing with hate speech, which includes both EU-wide instruments, such as Council Framework Decision 2008/913/JHA on combating certain forms of racism and xenophobia by means of criminal law (discussed below), and national laws (often criminal laws) against incitement to hatred.⁷⁹

Under the Convention, states enjoy a varying “margin of appreciation” to restrict hate speech. The margin of appreciation is a notion developed by the Strasbourg Court, which enables it to consider a country’s cultural, historical, and philosophical specificities.⁸⁰ This gives national authorities room for maneuver in applying the Convention.⁸¹ While the Strasbourg Court has generally been restrictive in the margin of appreciation granted to national authorities in relation to freedom of expression, it has accepted a wider margin with regard to issues such as hate speech.⁸²

Permissible speech restrictions must remain compatible with the right to freedom of expression. In general, when speech incites violence, hatred, or intolerance, it falls entirely outside the protection of Article 10 of the European Convention on Human Rights. Under the Convention, expression may be restricted, provided such limitations are prescribed by law, legitimate, and necessary in a democratic society. These restrictions may apply to both the creation and dissemination of such expression. The Strasbourg Court has consistently held that Article 10 protects not only inoffensive or neutral speech but also expression that may “offend, shock, or disturb.”⁸³ Yet its approach to hate speech — for which there is no clear definition or a consistent doctrinal framework — is not aligned with this core principle.⁸⁴ The Strasbourg Court currently risks undermining free speech by permitting restrictions on expression that is merely insulting or offensive, without requiring any element of incitement to violence or hostility.⁸⁵ The Future of Free Speech’s Senior Fellow Natalie Alkiviadou has criticized the Strasbourg Court’s jurisprudence on hate speech for its conceptual ambiguity, internal inconsistencies, and lack of empirical grounding.⁸⁶

One of the most problematic aspects of the Strasbourg Court’s case law is its inconsistent application of Article 17 of the Convention, commonly referred to as the “abuse clause.” Although this provision is meant to

⁷⁷ Regulation on a Single Market for Digital Services and Amending Directive 2000/31/EC (Digital Services Act), 2022 O.J. (L 2065) (October 19, 2022), <https://eur-lex.europa.eu/eli/reg/2022/2065/oj/eng>.

⁷⁸ European Commission, “Commission Opens Investigations to Safeguard Minors from Pornographic Content Under the Digital Services Act,” press release, May 27, 2025, <https://digital-strategy.ec.europa.eu/en/news/commission-opens-investigations-safeguard-minors-pornographic-content-under-digital-services-act>.

⁷⁹ “Responding to ‘Hate Speech’: Comparative Overview of Six EU Countries,” ARTICLE 19 (March 2018): 16, https://www.article19.org/wp-content/uploads/2018/03/ECA-hate-speech-compilation-report_March-2018.pdf.

⁸⁰ Thomson Reuters, Practical Law, “Margin of Appreciation,” accessed August 28, 2025, <https://uk.practicallaw.thomsonreuters.com/2-500-7514>.

⁸¹ Thomson Reuters, “Margin of Appreciation.”

⁸² Natalie Alkiviadou, *Hate Speech and the European Court of Human Rights* (Routledge, 2025), 96.

⁸³ *Handyside v. U.K.* (1976).

⁸⁴ Natalie Alkiviadou, “Hate Speech and the European Court of Human Rights: An Overview of My New Book,” Reason: The Volokh Conspiracy, July 15, 2025, <https://reason.com/volokh/2025/07/15/hate-speech-and-the-european-court-of-human-rights-an-overview-of-my-new-book>.

⁸⁵ Alkiviadou, “Overview of My New Book.”

⁸⁶ Alkiviadou, “Overview of My New Book.”

deny Article 10 protection to speech inciting violence or hatred, the court has often applied it on the basis of a speaker's ideology, effectively resulting in viewpoint discrimination.⁸⁷ The Strasbourg Court framework has resulted in a case law that is too lenient with vague national hate speech laws and that too hastily dismisses the chilling effects they create.⁸⁸ It is reasonable to assume that this framework, which excessively limits freedom of expression, will also be applied to AI-generated content.

A key instrument in the EU is the Council Framework Decision 2008/913/JHA, combating certain forms of racism and xenophobia by means of criminal law. This framework decision requires all member states to criminalize public incitement to violence or hatred against people due to their race, color, religion, descent, or national or ethnic origin. Countries can also criminalize incitement to violence or hatred based on other grounds. The majority of countries extend this to other protected characteristics, such as sexual orientation, and many also do so regarding gender identity.⁸⁹ The European Commission has called on several countries, such as Bulgaria, Ireland, Greece, and Hungary,⁹⁰ for failing to correctly transpose the framework decision. In the case of Ireland, for instance, the Commission considers that this country does not adequately govern incitement to hatred and violence, including the condoning, denial, or gross trivialization of international crimes and the Holocaust.⁹¹ Even below the threshold of criminal incitement, there are civil and administrative laws (and platform policies) addressing hateful expression.⁹²

A person who relies on generative AI to intentionally produce content to generate and then share punishable hate speech would presumably be held liable in the same way as if they had created it autonomously.

The liability of AI providers is more complex. Nevertheless, some avenues could lead to liability for AI companies. Notably, the European Commission has indicated that AI services may fall within the scope of the DSA, which could result in fines for noncompliance. At the time of writing, Meta was expected to submit a risk assessment regarding the AI features deployed across its platform.⁹³ Google has already submitted a risk assessment for its AI-generated search summaries.⁹⁴ Public information suggests that the Commission is considering whether ChatGPT should be subject to the DSA's systemic risk obligations.⁹⁵ Additionally, the Commission has indicated that DeepSeek's AI models could fall under the DSA if those models are integrated into other platforms.⁹⁶ The European Commission called in X to discuss its AI chatbot, Grok, following antisemitic comments Grok made.⁹⁷ Given the DSA's focus on combating hate speech and its obligations to limit the dissemination of illegal content, including illegal hate speech, on very large online platforms and search engines, AI providers could face significant fines. Importantly, when they generate new content, AI systems are highly unlikely to benefit from the "safe harbors" that grant some services immunity.⁹⁸

87 Alkiviadou, "Overview of My New Book."

88 Alkiviadou, "Overview of My New Book."

89 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.

90 See Representation in Ireland, "European Commission Calls on Ireland, Bulgaria and Estonia to Correctly Transpose EU Law Combating Racism and Xenophobia," October 3, 2024, https://ireland.representation.ec.europa.eu/news-and-events/news/european-commission-calls-ireland-bulgaria-and-estonia-correctly-transpose-eu-law-combating-racism-2024-10-03_en; European Commission, "January Infringements Package: Key Decisions," press release, January 26, 2023, https://ec.europa.eu/commission/presscorner/api/files/document/print/en/inf_23_142/INF_23_142_EN.pdf.

91 Representation in Ireland, "European Commission Calls on Ireland, Bulgaria and Estonia."

92 "Responding to 'Hate Speech,'" 23-26.

93 European Parliament, "Answer Given by Executive Vice-President Virkkunen."

94 Anna Ferrari, "Google's AI Overviews Face EU Compliance Review Under the DSA," *MLex*, August 8, 2025, <https://www.mlex.com/mlex/artificial-intelligence/articles/2375007>.

95 Luca Bertuzzi, "ChatGPT Faces Possible Designation as a Systemic Platform Under EU Digital Law," *MLex*, April 30, 2025, <https://www.mlex.com/mlex/artificial-intelligence/articles/2332484/chatgpt-faces-possible-designation-as-a-systemic-platform-under-eu-digital-law>.

96 European Parliament, "Answer Given by Vice-President Virkkunen."

97 Eliza Gkritsi, "EU Calls in X to Talk Grok After Antisemitic Outbursts," *Politico*, July 14, 2025, <https://www.politico.eu/article/european-commission-x-artificial-intelligence-chatbot-grok-antisemitism>.

98 Arcila, "Is It a Platform?," 460.

The AI Act does not explicitly address the generation of hate speech. However, it does require providers of powerful general-purpose AI models to assess and mitigate “systemic risks,” including potential negative effects on fundamental rights and “society as a whole.” The General-Purpose AI Code of Practice, which provides guidance on implementing the AI Act, identifies “hateful” and “radicalizing” content as risks that should be addressed and mitigated.⁹⁹ Although the code is not legally binding, it has been endorsed by major AI providers and carries significant influence. In this context, it is reasonable to expect that the European Commission may consider hate speech — and certainly illegal hate speech — to be a form of systemic risk that could trigger enforcement, including fines.

The systemic risk obligations under the AI Act and the DSA pose a significant concern for freedom of expression and access to information.¹⁰⁰ Their vagueness and broad scope leave them open to misuse by public authorities and may lead companies to engage in self-censorship. This phenomenon, already observed in the context of social media,¹⁰¹ appears to be repeating itself in the AI landscape.¹⁰² Implementing measures such as the Code of Conduct on Hate Speech¹⁰³ have often exacerbated the issue rather than addressed it, particularly by introducing strict deadlines for content removal.¹⁰⁴

It is doubtful whether these provisions, in their current state, are compatible with the principles of legality and proportionality established in Article 52 of the EU Charter of Fundamental Rights. Experts have pointed out that the systemic risk 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.”¹⁰⁵ They note a “fundamental tension” between the DSA’s risk-based approach and the legality principle.¹⁰⁶ Scholars also highlight concerns regarding proportionality, including that the assessment of proportionality is outsourced to companies and that proportionality is assessed in relation to the risks being mitigated rather than the legitimate aims pursued by the DSA.¹⁰⁷

2.6. Election and Political Content

The impact of generative AI on politics and elections is now a mainstream concern in Europe. AI tools present new challenges, as they can be used to fabricate news, impersonate candidates, or disseminate propaganda. Still, in democratic societies, political speech is highly protected: Citizens must be free to criticize government and advocate for change.

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

100 Jordi Calvet-Bademunt, “Who Decides What’s Good for Society? AI, the DSA, and the Future of Expression in Europe,” *The Bedrock Principle*, June 12, 2025, <https://www.bedrockprinciple.com/p/who-decides-whats-good-for-society>.

101 “Preventing ‘Torrents of Hate’ or Stifling Free Expression Online?,” *The Future of Free Speech*, May 28, 2024, <https://futurefreespeech.org/preventing-torrents-of-hate-or-stifling-free-expression-online>.

102 Jordi Calvet-Bademunt, Jacob Mchangama, and Isabelle Anzabi, “One Year Later: AI Chatbots Show Progress on Free Speech — But Some Concerns Remain,” *The Bedrock Principle*, April 1, 2025, <https://www.bedrockprinciple.com/p/one-year-later-ai-chatbots-show-progress>; Jordi Calvet-Bademunt and Jacob Mchangama, “Freedom of Expression in Generative AI: A Snapshot of Content Policies,” *The Future of Free Speech*, February 2024, https://futurefreespeech.org/wp-content/uploads/2023/12/FFS_AI-Policies_Formatting.pdf.

103 The DSA Code of Conduct on Countering Illegal Hate Speech Online +, adopted in 2025, builds on the Code of Conduct adopted in 2016. See European Commission, “Code of Conduct on Countering Illegal Hate Speech Online +,” policy publication, January 20, 2025, <https://digital-strategy.ec.europa.eu/en/library/code-conduct-countering-illegal-hate-speech-online>.

104 Natalie Alkiviadou, “Why We Should Be Concerned About the ‘Illegal Hate Speech’ Code of Conduct +,” *The Bedrock Principle*, February 5, 2025, <https://www.bedrockprinciple.com/p/why-we-should-be-concerned-about>.

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

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

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

The fear is that AI-generated fake content — such as deepfake videos of politicians — could deceive voters or sow chaos. Throughout 2023 and 2024, prominent media outlets expressed concerns about the potential influence of AI on elections.¹⁰⁸ A vice president of the European Commission described deepfakes of politicians as “an atomic bomb [that could] change the course of voter preferences.”¹⁰⁹ Public concern matched the media warnings. Approximately 40% of European voters expressed concerns about the misuse of AI during elections.¹¹⁰ Yet research suggests the fear-driven narrative about AI in 2024 was not supported by evidence.¹¹¹ In the words of two leading researchers who analyzed every instance of AI use in elections collected in the WIRED AI Elections Project, “misinformation is not a problem.”¹¹² These findings also reflect the limited effect of “disinformation” more broadly.¹¹³ Still, these fears have led to overly restrictive or vague legislation.¹¹⁴

The EU is addressing the challenges of generative AI with a combination of legal requirements. Some of these requirements rely on transparency, such as labeling deepfakes, and others are more restrictive, notably the obligations on systemic risk imposed by the DSA and the AI Act. These efforts are complemented by legislation specifically targeting disinformation, which raises significant concerns for freedom of expression.

Part of the EU’s approach has been to mandate transparency for AI-generated content. Under the EU AI Act, providers and deployers of AI systems that generate or manipulate content — such as deepfakes — are required to ensure that such content is marked or labeled to disclose its artificial origin.¹¹⁵ This labeling might involve visible indicators, such as watermarks, or metadata tags, that are machine-readable and detectable as artificially generated or manipulated. The act includes specific exceptions to these obligations. Transparency requirements are relaxed for artistic, creative, satirical, fictional, or similar works, such as AI-generated movies or parodies. In these cases, disclosure of AI involvement should be done in a manner that does not disrupt the viewer’s experience. In addition, if AI-generated text content informing users on matters of public interest has undergone human review or editorial control, with a natural or legal person holding editorial responsibility for the publication, labeling requirements may be relaxed. The effectiveness of watermarking and labeling and its impact on freedom of expression are currently under discussion.¹¹⁶

In addition to requiring transparency, the AI Act’s obligations on systemic risk, mentioned above, may limit the content that can be generated. Under the AI Act, providers of general-purpose AI models with systemic risk are required to mitigate such risk. This is also the case under the DSA for VLOPs and VLOSEs, which may be powered by generative AI. The same dangers we addressed regarding hate speech — resulting from the vagueness and breadth of the provisions — apply here. While most member states of the EU are robust democracies operating under the rule of law, and the EU is institutionally committed to these values, these obligations may also be misused because of their vagueness and breadth.¹¹⁷ Indeed, former commissioner

108 See, inter alia, Pranshu Verma and Cat Zakrzewski, “AI Deepfakes Threaten to Upend Global Elections: No One Can Stop Them,” *Washington Post*, April 23, 2024, <https://www.washingtonpost.com/technology/2024/04/23/ai-deepfake-election-2024-us-india/>; Lorne Cook and Kelvin Chan, “AI Could Supercharge Disinformation and Disrupt EU Elections, Experts Warn,” *Associated Press*, June 5, 2024, <https://apnews.com/article/eu-european-union-election-disinformation-43b7e4017825d9d382859894b7625e7a>.

109 Henry Foy, “Why Big Tech and Deepfakes Keep EU Election Guardians Up at Night,” *Financial Times*, February 28, 2024, <https://www.ft.com/content/e9f4d2c0-5d33-409f-8e60-ed9eac7febb>.

110 “European Tech Insights 2024,” Center for the Governance of Change, IE University (2024): 5, <https://static.ie.edu/CGC/European%20Tech%20Insights%202024%20-%20IE%20CGC.pdf>.

111 Sayash Kapoor and Arvind Narayanan, “We Looked at 78 Election Deepfakes: Political Misinformation Is Not an AI Problem,” Knight First Amendment Institute, Columbia University, December 13, 2024, <https://knightcolumbia.org/blog/we-looked-at-78-election-deepfakes-political-misinformation-is-not-an-ai-problem>; Sam Stockwell et al., “AI-Enabled Influence Operations: Safeguarding Future Elections,” Center for Emerging Technology and Security, Alan Turing Institute, November 13, 2024, <https://cetas.turing.ac.uk/publications/ai-enabled-influence-operations-safeguarding-future-elections>; Sam Stockwell, “AI-Enabled Influence Operations: Threat Analysis of the 2024 UK and European Elections,” Center for Emerging Technology and Security, Alan Turing Institute, September 19, 2024, <https://cetas.turing.ac.uk/publications/ai-enabled-influence-operations-threat-analysis-2024-uk-and-european-elections>.

112 Kapoor and Narayanan, “We Looked at 78 Election Deepfakes.”

113 Alexander Hohlfeld, “Feedback to the European Commission on the ‘European Democracy Shield’: Call for Evidence,” *The Future of Free Speech*, May 23, 2025, <https://futurefreespeech.org/wp-content/uploads/2025/05/EU-Democracy-Shield-Response.pdf>.

114 Jordi Calvet-Bademunt, “The AI Election Panic: How Fear-Driven Policies Could Limit Free Expression,” *Tech Policy Press*, April 2, 2025, <https://www.techpolicy.press/the-ai-election-panic-how-fear-driven-policies-could-limit-free-expression>.

115 Under the AI Act, a deployer is any natural or legal person using an AI system, except where the AI system is used in the course of a personal non-professional activity.

116 Scott Babwah Brennen, “Will AI Content Labels Work?,” Center on Technology Policy, New York University, October 8, 2024, https://techpolicy.nyu.org/wp-content/uploads/2024/10/CTP_Will-AI-content-labels-work_final.pdf.

117 Calvet-Bademunt, “Who Decides What’s Good for Society?”

Thierry Breton undertook several controversial informal actions under the DSA, which contains similar provisions. These included sending letters to companies that conflated illegal content with disinformation,¹¹⁸ as well as warning X about its DSA systemic risk obligations and cautioning Elon Musk against “harmful content” ahead of an interview between Musk and then-candidate Donald Trump on the platform.¹¹⁹ More recently, Poland’s government called on the European Commission to investigate possible violations of EU law by Grok, xAI’s chatbot.¹²⁰ The chatbot generated a series of attacks on prominent European politicians, Poland’s president for one, and antisemitic and hateful outbursts.¹²¹

The GPAI Code of Practice, although not mandatory, is significantly influential in the implementation of the AI Act and has the backing of major companies, yet it does little to alleviate these concerns. It identifies “false” and “radicalizing” content as risks that companies should consider mitigating.¹²² These obligations raise concerns regarding misguided enforcement, similar to those in the DSA, and self-censorship.¹²³

The European Commission published guidelines for providers of VLOPs and VLOSEs to mitigate systemic risks to electoral processes under the DSA. In this document, the Commission recommends companies adapt their terms and conditions “to significantly decrease the reach and impact of generative AI content that depicts disinformation or misinformation.”¹²⁴ Given the contentious nature of disinformation, this recommendation could pose a threat to freedom of expression. The guidelines also encourage companies to label deepfakes clearly.¹²⁵ The European Commission has endorsed a Code of Conduct on Disinformation as well, signed by major tech companies; however, this code does not explicitly address generative AI.

Beyond the systemic risk obligations, the AI Act designates AI systems “intended to be used for influencing the outcome of an election or referendum or the voting behaviour” as high-risk. Such systems must implement a risk management framework to identify and address potential risks to health, safety, or fundamental rights. If applied in a manner that restricts certain voices but not others, this obligation could raise concerns about freedom of expression. How it will be implemented in practice remains to be seen.

In addition to EU-wide measures, some member states have passed their own laws against disinformation. In 2018, France adopted a law that empowered judges to order the removal of “fake news” within 48 hours during election campaigns.¹²⁶ Several countries, such as Croatia, the Czech Republic, and Hungary, have criminal laws that cover the deliberate dissemination of disinformation posing a threat to peace or public order.¹²⁷ These laws raise significant concerns regarding freedom of expression. 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” and susceptible to providing authorities with “excessive discretion to determine what is disinformation, what is a mistake, what is truth.”¹²⁸

118 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>.

119 Thierry Breton (@ThierryBreton), “With great audience comes greater responsibility #DSA As there is a risk of amplification of potentially ...,” X, <https://x.com/ThierryBreton/status/1823033048109367549/photo/1>.

120 Scott Roxborough, “Poland Slams Grok AI for Antisemitic Outbursts, Urges EU Probe,” *Hollywood Reporter*, July 10, 2025, <https://www.hollywoodreporter.com/news/politics-news/grok-ai-poland-eu-probe-antisemitic-outbursts-1236311137>.

121 Roxborough, “Poland Slams Grok AI.”

122 European Commission, “Safety and Security,” 34.

123 Calvet-Bademunt, “Who Decides What’s Good for Society?”

124 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>.

125 Guidelines for Providers of Very Large Online Platforms, para. 40.

126 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 [J.O.], December 22, 2018, <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000037847559> (Fr.).

127 Konrad Bleyer-Simon, Elda Brogi, Iva Nenadić, and Teona Nesović, “Policies to Tackle Disinformation in EU Member States — Part II,” European Digital Media Observatory, 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>.

128 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

In *Salov v. Ukraine* (2005), the Strasbourg Court held that a prosecution for “dissemination of false information” under Ukraine’s election legislation violated the right to freedom of expression. Notably, the court held that the Convention “does not prohibit discussion or dissemination of information received even if it is strongly suspected that this information might not be truthful.”¹²⁹ This precedent was also followed in cases against Poland involving the prohibition of false news and untrue information.¹³⁰ The court’s case law suggests reluctance to uphold restrictions that do not distinguish deliberate lies from unfounded but good-faith statements.¹³¹ In this regard, the Strasbourg Court found no violation of Article 10 of the Convention in Poland’s application of its electoral disinformation law in *Staniszewski v. Poland* (2021).¹³² The applicant, a journalist, had alleged that a local mayor selected a village for a regional harvest festival solely to boost electoral support there. Crucially, he was given ample opportunity to substantiate or at least attempt to verify this claim. While the “untrue” allegation had to be assessed in light of his role as a political journalist, he remained, according to the court, under a duty to take good-faith steps to verify factual allegations directed at an election candidate.¹³³

Although not acting under its fake news law, France was the first EU member state to shut down a platform — TikTok — following riots in New Caledonia, arguing that TikTok had been used to spread disinformation promoted by Azerbaijan.¹³⁴ Romania went so far as to annul its presidential elections in 2024, alleging Russian interference,¹³⁵ and the European Commission subsequently opened formal proceedings under the DSA for potentially failing to mitigate the risks to elections.¹³⁶

In line with the findings of the 2018 High Level Expert Group on Fake News and Online Disinformation, established by the European Commission, the EU and its member states should avoid “any form of censorship, either public or private.”¹³⁷ They should instead rely on media literacy, empower users and journalists to tackle disinformation, safeguard the diversity and sustainability of the European news media ecosystem, promote research, and enhance transparency of online news.¹³⁸

2.7. Copyright

2.7.1. Generative AI Inputs

The first key question to address in the context of AI and copyright law is how AI training is governed. AI models are developed using deep learning methods, which adjust the models’ internal parameters based on training data. Building these datasets begins with data collection, which often draws heavily on freely accessible online resources. Model providers typically treat the composition of their training datasets as confidential, considering them commercially sensitive. In contrast, many rights holders worry about losing control over their works, and researchers have documented multiple ongoing lawsuits outside the EU alleging that GPAI training data includes copyrighted material.¹³⁹

Alkiviadou, “Prison for Fake News,” *Verfassungsblog*, July 12, 2024, <https://verfassungsblog.de/prison-for-fake-news>.

¹²⁹ *Salov v. Ukraine*, App. No. 65518/01, Eur. Ct. H.R. (September 6, 2005), <https://hudoc.echr.coe.int/eng/?i=001-70096>, discussed in Alkiviadou, “Prison for Fake News.”

¹³⁰ Alkiviadou, “Prison for Fake News.”

¹³¹ Ethan Shattock, “Fake News in Strasbourg: Electoral Disinformation and Freedom of Expression in the European Court of Human Rights (ECtHR),” *European Journal of Law and Technology* 13, no. 1 (2022): 20, <https://ejlt.org/index.php/ejlt/article/view/882>.

¹³² *Staniszewski v. Poland*, App. No. 20422/15, Eur. Ct. H.R. (October 14, 2021), <https://hudoc.echr.coe.int/fre/?i=001-212158>, discussed in Shattock, “Fake News in Strasbourg,” 20.

¹³³ Shattock, “Fake News in Strasbourg,” 20.

¹³⁴ Paul Kirby, “TikTok Ban Lifted as New Caledonia Emergency Ends,” *BBC News*, May 29, 2024, <https://www.bbc.com/news/articles/c0dd94jv9jpo>; “TikTok Censorship in New Caledonia: A Review of a Democratic Failure,” *La Quadrature du Net*, June 6, 2024, <https://www.laquadrature.net/en/2024/06/06/tiktok-censorship-in-new-caledonia-a-review-of-a-democratic-failure>.

¹³⁵ Atlantic Council Experts, “Romania Annulled Its Presidential Election Results Amid Alleged Russian Interference: What Happens Next?,” *New Atlanticist*, December 6, 2024, <https://www.atlanticcouncil.org/blogs/new-atlanticist/romania-annulled-its-presidential-election-results-amid-alleged-russian-interference-what-happens-next>.

¹³⁶ European Commission, “Commission Opens Formal Proceedings Against TikTok on Election Risks Under the Digital Services Act,” press release, December 16, 2024, https://ec.europa.eu/commission/presscorner/detail/en/ip_24_6487.

¹³⁷ 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>.

¹³⁸ European Commission, “Final Report of the High Level Expert Group.”

¹³⁹ Tristan Marcelin 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>.

To gather publicly available online data, developers commonly employ web crawlers — automated programs designed to browse and perform specified tasks systematically across the internet.¹⁴⁰

Copyright law grants authors exclusive economic and moral rights, including the rights to reproduce, distribute, and communicate works and to make works available to the public.¹⁴¹ The EU Copyright Directive introduced two specific exceptions to these rights for text and data mining (TDM).¹⁴² TDM refers to the use of automated techniques to analyze text and data to generate information, such as patterns, trends, and correlations.¹⁴³ The directive allows anyone, including commercial entities, to mine lawfully accessible content for any purpose, provided the rights holder has not expressly reserved their rights.¹⁴⁴ It also establishes a more substantial exemption for research organizations and cultural heritage institutions, enabling them to carry out TDM for scientific research without giving rights holders the possibility to opt out.¹⁴⁵

In practice, many AI companies have relied on the TDM exception to build their training datasets. However, this has led to pushback from authors, publishers, and other content creators.¹⁴⁶ Some scholars have noted that, although TDM may seem synonymous with AI training, AI training extends beyond TDM.¹⁴⁷ Others consider TDM to cover AI training.¹⁴⁸ More broadly, most seem to agree that EU law does not comprehensively address the intellectual property issues raised by this aspect of AI.¹⁴⁹ Some member states have stated that copyright uses for AI training go beyond the scope of the TDM exception.¹⁵⁰ The Luxembourg Court is currently considering whether the TDM exception applies to AI training.¹⁵¹

The EU Information Society Directive (Directive 2001/29) also provides an exception for temporary reproductions made in the course of technological processes. However, it is widely accepted that this defense is unlikely to apply to AI training, in part because the copies produced must lack independent economic value — a condition that AI training often does not meet.¹⁵² Hence, AI providers cannot rely on this exemption to train their models.

The EU AI Act contains two provisions that are particularly relevant. First, it requires GPAI providers to comply with copyright law and the Copyright Directive's opt-out rule, which permits TDM unless rights holders explicitly object. Second, it obliges providers to publish a sufficiently detailed summary of the training data content. To aid implementation, the European Commission included a dedicated chapter on copyright in its GPAI Code of Practice. The Commission also published a template for GPAI model providers to summarize the data used to train their models.

training-of-general%E2%80%91purpose-ai.

140 Marcelin and Cassetti, "AI and Copyright."

141 Marcelin and Cassetti, "AI and Copyright."

142 Marcelin and Cassetti, "AI and Copyright."

143 Directive 2019/790 on Copyright and Related Rights in the Digital Single Market and Amending Directives 96/9/EC and 2001/29/EC, 2019 O.J. (L 130/92), art. 2, para. 2 (May 17, 2019), <https://eur-lex.europa.eu/eli/dir/2019/790/oj/eng>.

144 Directive 2019/790 on Copyright and Related Rights, 2019 O.J. (L130), art. 4.

145 Directive 2019/790 on Copyright and Related Rights, 2019 O.J. (L130), art. 3. In a noteworthy ruling, the Hamburg Regional Court suggested that the notion of a "research organization" could be interpreted broadly — too broadly, in the view of some scholars. Eleonora Rosati, "Copyright Exceptions and Fair Use Defences for AI Training Done for 'Research' and 'Learning,' or the Inescapable Licensing Horizon," *European Journal of Risk Regulation*, July 29, 2025: 4, <https://doi.org/10.1017/err.2025.10035>.

146 Jennifer Rankin, "EU Accused of Leaving 'Devastating' Copyright Loophole in AI Act," *The Guardian*, February 19, 2025, <https://www.theguardian.com/technology/2025/feb/19/eu-accused-of-leaving-devastating-copyright-loophole-in-ai-act>.

147 Rosati, "Copyright Exceptions and Fair Use Defences," 4; Tim W. Dornis and Sebastian Stober, "Generative AI Training and Copyright Law," *Transactions of the International Society for Music Information Retrieval* (2025), <https://doi.org/10.48550/arXiv.2502.15858>.

148 Philipp Hacker, "Copyright, AI, and the Future of Internet Search Before the CJEU," *Verfassungsblog*, July 17, 2025, <https://verfassungsblog.de/copyright-ai-cjeu>.

149 See Marcelin and Cassetti, "AI and Copyright"; Hacker, "Copyright, AI, and the Future of Internet Search."

150 Marcelin and Cassetti, "AI and Copyright."

151 Hacker, "Copyright, AI, and the Future of Internet Search."

152 Rosati, "Copyright Exceptions and Fair Use Defences," 5.

2.7.2. Generative AI Outputs

Another concern is how copyright law applies to AI outputs, with two critical aspects: first, the relationship between AI outputs and the materials used to train the model; second, whether AI-generated content itself can be protected by copyright.¹⁵³

Regarding the first issue, in general, an AI output can infringe legal rights in two primary ways. If it contains substantial, direct similarities to legally protected elements of an existing work, it would likely breach that work's reproduction rights.¹⁵⁴ And if protected aspects of an existing work are incorporated into the output through unauthorized adaptations or modifications, the result would likely be considered a derivative work based on the original.¹⁵⁵

A key question is whether, and to what extent, copyright exceptions apply to AI-generated content.¹⁵⁶ These could include exceptions for quotation, criticism, or review or for uses such as caricature, parody, or pastiche.¹⁵⁷

If a violation is found, the person who prompted the AI model would likely be held primarily responsible, as they are the one who brings the infringing content into existence.¹⁵⁸ However, AI providers and developers may also face liability if they fail to take reasonable precautions.¹⁵⁹ Recent rulings by the Luxembourg Court have established that platforms must fulfill specific duties of care to avoid liability for copyright violations by users. Though these duties are not directly applicable to AI models, it is possible that they are transposed to this context.¹⁶⁰ While many AI providers' terms of service disclaim liability, the court has made clear that such clauses, on their own, do not shield providers from responsibility.¹⁶¹

The second key issue is whether AI-created content is itself protected by copyright. Under EU law, copyright requires an "original work" that is the author's own intellectual creation — traditionally interpreted to require human creativity.¹⁶² If the output is primarily the result of human effort, created with the assistance of AI, it is almost certainly protected by copyright.¹⁶³ If, by contrast, the human involvement in the generation of content is minimal — for instance, simply providing a prompt — it would not qualify for copyright at all.¹⁶⁴

2.8. Measures Empowering Freedom of Expression

The debate around generative AI often centers on potential risks and restrictions, which is why it is crucial to highlight how generative AI can strengthen freedom of expression — particularly for linguistic and cultural minorities, not to mention society at large, through education and creative opportunities. EU institutions are working to harness AI as a catalyst for more inclusive and diverse expression.

¹⁵³ Novelli et al., "Generative AI in EU Law," 10.

¹⁵⁴ Novelli et al., "Generative AI in EU Law," 11.

¹⁵⁵ Novelli et al., "Generative AI in EU Law," 11.

¹⁵⁶ Novelli et al., "Generative AI in EU Law," 11; Eleonora Rosati, "Infringing AI: Liability for AI-Generated Outputs Under International, EU, and UK Copyright Law," *European Journal of Risk Regulation* 16 (2025): 621–25, <https://doi.org/10.1017/err.2024.72>.

¹⁵⁷ João Pedro Quintais, "Generative AI, Copyright and the AI Act," *Computer Law & Security Review* 56, no. 106107 (2025): 4, <https://doi.org/10.1016/j.clsr.2025.106107>.

¹⁵⁸ Rosati, "Infringing AI," 619; Novelli et al., "Generative AI in EU Law," 11.

¹⁵⁹ Rosati, "Infringing AI," 619; Novelli et al., "Generative AI in EU Law," 11.

¹⁶⁰ Novelli et al., "Generative AI in EU Law," 11.

¹⁶¹ Rosati, "Infringing AI," 621.

¹⁶² Novelli et al., "Generative AI in EU Law," 11.

¹⁶³ Novelli et al., "Generative AI in EU Law," 11.

¹⁶⁴ Novelli et al., "Generative AI in EU Law," 11.

2.8.1. Linguistic and Cultural Diversity

The EU has 24 official languages and about 60 regional and minority languages.¹⁶⁵ The EU Charter of Fundamental Rights and the Treaty on the European Union both highlight the importance of protecting cultural and linguistic diversity.¹⁶⁶

Experts have noted that major AI models are predominantly trained on English-language and other high-resource language data.¹⁶⁷ This causes models to underperform in low-resource languages.¹⁶⁸ While the resource gap is particularly troubling in low- and middle-income countries, it also exists in Europe with languages such as Welsh and Basque.¹⁶⁹

The EU has launched two important initiatives to address the shortage of European language data needed for training large language models (LLMs) and to support cultural diversity — the Alliance for Language Technologies European Digital Infrastructure Consortium (ALT-EDIC) and the Language Data Space (LDS). These initiatives are designed to break down language barriers in the EU and provide more accessible solutions for smaller businesses, helping preserve the EU’s cultural and linguistic heritage and strengthening Europe’s quest for tech sovereignty. Formed in February 2024, ALT-EDIC includes 17 participating member states and nine observer member states and regions. LDS aims to create a cohesive marketplace for language data and to enhance the collection and sharing of multilingual data that will support the development of European LLMs.¹⁷⁰ The LLMs4EU project, a recent initiative launched by ALT-EDIC, aims to preserve European linguistic and cultural diversity by ensuring that LLMs and the tools needed to use them in all EU languages are available as open data.¹⁷¹

In addition, the GenAI4EU program supports start-ups and small and medium-sized enterprises in developing, adapting, and validating generative AI models and solutions across multiple sectors, including the cultural and creative sectors and industries. In this area, GenAI4EU calls for generative AI systems that cater to cultural and linguistic diversity.¹⁷²

Finally, an alliance of universities, companies, and other stakeholders has launched a pan-European initiative aimed at creating a family of open AI models covering all EU official languages.¹⁷³

2.8.2. Media and AI Literacy

Media literacy — the ability to access, analyze, and engage with information effectively and safely¹⁷⁴ — constitutes a powerful alternative to restrictions on freedom of expression in addressing challenges such as

¹⁶⁵ European Union, “Languages,” accessed August 29, 2025, https://european-union.europa.eu/principles-countries-history/languages_en; Directorate-General for Education, Youth, Sport and Culture, “Linguistic Diversity in the European Union: Examples of Projects Supporting Regional and Minority Languages,” European Commission (2024), <https://op.europa.eu/en/publication-detail/-/publication/d325c589-011a-11ef-a251-01aa75ed71a1>.

¹⁶⁶ Charter of Fundamental Rights of the E.U., art. 22; Treaty on European Union, art. 3.

¹⁶⁷ “How AI Is Leaving Non-English Speakers Behind,” *Stanford Report*, May 19, 2025, <https://news.stanford.edu/stories/2025/05/digital-divide-ai-llms-exclusion-non-english-speakers-research>.

¹⁶⁸ “How AI Is Leaving Non-English Speakers Behind.”

¹⁶⁹ Juan N. Pava, Caroline Meinhardt, Haifa Badi Uz Zaman, Toni Friedman, Sang T. Truong, Daniel Zhang, Elena Cryst, Vukosi Marivate, and Sanmi Koyejo, “Mind the (Language) Gap: Mapping the Challenges of LLM Development in Low-Resource Language Contexts,” Asia Foundation and Stanford Institute for Human-Centered Artificial Intelligence white paper (2025): 8, <https://hai-production.s3.amazonaws.com/files/hai-taf-pretoria-white-paper-mind-the-language-gap.pdf>.

¹⁷⁰ 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>.

¹⁷¹ “LLMs4EU,” OpenAIRE, accessed September 10, 2025, <https://www.openaire.eu/llms4eu>.

¹⁷² “GenAI4EU: Creating European Champions in Generative AI,” European Innovation Council, accessed September 10, 2025, https://eic.ec.europa.eu/eic-funding-opportunities/eic-accelerator/eic-accelerator-challenges-2025/genai4eu-creating-european-champions-generative-ai_en.

¹⁷³ “OpenEuroLLM — A European Family of Large Language Models,” European Union, accessed August 29, 2025, https://strategic-technologies.europa.eu/be-inspired/step-stories/openeurolm-european-family-large-language-models_en.

¹⁷⁴ Tarja Laaninen, “Media Literacy: Fostering a Key Civic Skill in a Digital Information Environment,” EPRS Briefing PE 772.886 (May 2025): 1, [https://www.europarl.europa.eu/RegData/etudes/BRIE/2025/772886/EPRS_BRI\(2025\)772886_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2025/772886/EPRS_BRI(2025)772886_EN.pdf).

disinformation. As noted by the EU High Level Expert Group on Fake News and Online Disinformation, media literacy is a key pillar of a robust strategy against disinformation.¹⁷⁵ It is a critical skill that empowers citizens to navigate today's complex news environment and make informed decisions.¹⁷⁶ In the AI context, strong media literacy enables people to critically assess and responsibly use AI-generated content, thereby strengthening democratic debate while mitigating the risks of manipulation, censorship, and diminished free expression. Many of the EU's efforts focus on media literacy in general.

The Audiovisual Media Services Directive requires member states to promote media literacy skills.¹⁷⁷ The DSA Code of Practice on Disinformation, which supports implementation of the DSA, commits signatories to maintain and strengthen their efforts in media literacy and critical thinking.¹⁷⁸ The Code of Conduct on Countering Illegal Hate Speech Online also highlights the role of critical thinking and encourages its signatories to promote media literacy.¹⁷⁹ In its guidelines for providers of very large online platforms and very large online search engines on mitigating systemic risks to electoral processes, the European Commission likewise underscores the importance of media literacy initiatives.¹⁸⁰

In her 2024–2029 political guidelines, European Commission president Ursula von der Leyen emphasized the need for digital and media literacy.¹⁸¹ A forthcoming initiative in this area is the European Democracy Shield, which aims to strengthen democratic societies against manipulation. It is fundamental that the new shield, for which the European Commission issued a call for evidence in the spring of 2025, is based on sound research and promotes democratic discourse, viewpoint diversity, and competing values rather than censorship.¹⁸²

The European Commission also supports the European Digital Media Observatory (EDMO), which unites fact-checkers, media literacy experts, and researchers to combat disinformation. EDMO identifies best practices, fosters knowledge exchange across Europe, and runs awareness campaigns. EDMO is specifically exploring the impact of AI on information integrity.¹⁸³

In addition, the EU, in partnership with the OECD, is developing the AI Literacy Framework (AILit Framework) to empower primary and secondary school students with the critical-thinking skills needed to understand, engage with, and innovate using digital technologies.¹⁸⁴

Finally, under the AI Act, providers and deployers of AI systems are required to ensure that their staff, as well as any other individuals involved in operating or using these systems on their behalf, possess an adequate level of AI literacy. According to the AI Act, AI literacy refers to the skills, knowledge, and understanding that enable providers, deployers, and affected individuals to make informed decisions about the deployment of AI systems and to develop awareness of both the opportunities and the risks of AI, including potential harm it may cause.¹⁸⁵

¹⁷⁵ European Commission, "Final Report of the High Level Expert Group."

¹⁷⁶ Laaninen, "Media Literacy," 1.

¹⁷⁷ European Commission, "Media Literacy," updated October 15, 2024, <https://digital-strategy.ec.europa.eu/en/policies/media-literacy>.

¹⁷⁸ European Commission, "Code of Conduct on Disinformation," policy publication, February 13, 2025, <https://digital-strategy.ec.europa.eu/en/library/code-conduct-disinformation>.

¹⁷⁹ European Commission, "Code of Conduct on Countering Illegal Hate Speech Online +," para. 5.1 and app. 2, para. 5.

¹⁸⁰ Guidelines for Providers of Very Large Online Platforms, para. 40.

¹⁸¹ Laaninen, "Media Literacy," 1.

¹⁸² 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-distrusting-its-citizens>.

¹⁸³ European Digital Media Observatory (EDMO), "Artificial Intelligence," accessed August 29, 2025, <https://edmo.eu/thematic-areas/artificial-intelligence>.

¹⁸⁴ AILit Framework, accessed August 29, 2025, <https://ailiteracyframework.org>.

¹⁸⁵ European Commission, "AI Literacy — Questions & Answers," updated August 18, 2025, <https://digital-strategy.ec.europa.eu/en/faqs/ai-literacy-questions-answers>.

2.8.3. Accessibility in AI

The EU AI Act hardwires accessibility into the life cycle of high-risk AI systems: Providers must ensure compliance — by design — with EU accessibility law (the Web Accessibility Directive and the European Accessibility Act). Beyond design, when AI interacts with people or outputs synthetic content, providers and deployers face transparency duties (e.g., notifying users, marking outputs as AI-generated). These notices must meet applicable accessibility requirements so that persons with disabilities can use and understand them. Finally, the act bans exploitative uses that prey on vulnerabilities linked to age or disability when they materially distort behavior and risk significant harm, drawing a clear line against manipulative or discriminatory practices.

3. Conclusion

Generative AI has the potential to expand access to information and enable new forms of creativity and participation in public debate. Yet, as we show in this chapter, some of the new European rules risk narrowing the space for lawful expression rather than enlarging it.

The EU's legal protections for speech, grounded in Article 10 of the Convention and Article 11 of the Charter, set a clear standard in principle: Expression is protected even when it offends, shocks, or disturbs.¹⁸⁶ However, the Strasbourg Court has been too willing to uphold restrictions, particularly in cases involving hate speech, applying a deferential margin of appreciation that leaves a wide scope for overbroad national laws. This permissive stance, now intersecting with vague, far-reaching AI-related rules, risks excessively limiting the voices and perspectives citizens have access to.

The AI Act, the DSA, and other similar instruments contain measures that can serve legitimate purposes when narrowly drawn — such as prohibiting the most harmful forms of content, including AI-generated CSAM — and recognizing exemptions for artistic, creative, satirical, or fictional works. But a key part of the EU's new AI regime — i.e., systemic risk obligations — remains deeply problematic. The concept of “systemic risks” is left so undefined that it could encompass a wide range of controversial or unpopular ideas. Overlapping compliance obligations — particularly for general-purpose AI models under the DSA — create strong incentives for providers to over-remove content, especially in politically sensitive contexts. In practice, the combination of legislative vagueness, regulatory discretion, and the Strasbourg Court's current leniency toward restrictions could make AI-assisted political speech, journalism, or satire vulnerable to suppression.

These trends point to a structural issue: Rather than applying the whole discipline of necessity and proportionality tests before restricting speech, EU and Council of Europe actors risk building a framework where lawful expression is filtered out preemptively, not because it violates the law but because it is safer for intermediaries to take it down to avoid legal and reputational risks. That logic, once embedded in AI governance, will be difficult to reverse.

Europe must decide whether to approach generative AI as a tool for broadening the marketplace of ideas or as a force to be controlled within narrow, risk-averse limits. The critical question for policymakers, regulators, and courts is this: Will the AI future be one in which technological innovation strengthens freedom of expression, or one in which it narrows the scope of permissible speech?

¹⁸⁶ *Handyside v. U.K.* (1976), para. 49.



**THE
FUTURE
OF
FREE
SPEECH**

OCTOBER 2025

www.futurefreespeech.org