

THE FUTURE OF FREE SPEECH

THAT VIOLATES MY POLICIES 2026

AI LAWS, CHATBOTS, AND THE FUTURE OF EXPRESSION

DIRECTED BY:
Jordi Calvet-Bademunt,
Jacob Mchangama,
and Isabelle Anzabi

www.futurefreespeech.org

ACKNOWLEDGMENTS

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

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

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

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

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

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







© The Future of Free Speech 2026
Publication Date: October 2026

PREFACE

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

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

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

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

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

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

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

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

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



ARTIFICIAL INTELLIGENCE & FREEDOM OF EXPRESSION IN THE UNITED KINGDOM

By
Eliza Bechtold



ELIZA BECHTOLD

Dr. Eliza Bechtold is the Programmes Manager and a Research Fellow at the Bonavero Institute of Human Rights at the University of Oxford, where she is also a member of the Law Faculty. She researches in the area of freedom of expression, in particular, the regulation of extreme speech in the United States and Europe. She holds a PhD from Durham Law School, an LLM from the London School of Economics and Political Science, a JD from UC Law San Francisco (formerly UC Hastings) and a BA in Legal Studies from Scripps College. Prior to entering academia, Dr. Bechtold worked as an attorney in the United States, serving as Legal Director of a state affiliate of the American Civil Liberties Union and as a litigation associate at law firms, including DLA Piper LLP.

ABSTRACT

While the Labour government has promised to pass legislation specifically directed to regulating AI, at present, the UK does not have legislation covering AI as a technology, and there is no single AI regulator in the country. Instead, AI is regulated according to how it is used and to the extent such use falls within the scope of existing statutory frameworks. Additionally, there is an extensive collection of policy initiatives that directly target AI and the opportunities it presents in terms of the economy, law enforcement, and other sectors. These developments are set against a backdrop of protections for freedom of expression at the national and regional levels. As a Council of Europe member state, the UK is subject to the jurisdiction of the European Court of Human Rights and has obligations to comply with the European Convention on Human Rights. The latter is incorporated into the domestic legal system by way of the Human Rights Act 1998. Notwithstanding these protections, the government's enthusiasm to leverage AI in the public and private sectors as well as its regulation by way of legislation such as the Online Safety Act 2023, raises concerns that there is insufficient attention being paid to the potential impact on freedom of expression and other human rights in this context. Additionally, while the government is investing billions of pounds in AI-related initiatives, there is no suggestion that it intends to increase funding for the independent statutory body responsible for protecting the human rights of those living in the UK. Collectively, these developments suggest that the UK is not taking a human rights-based approach to AI regulation and that freedom of expression is unlikely to be adequately protected as the government moves forward with its ambitious plans to shape what it characterizes as an AI revolution.

1. INTRODUCTION

The evolving landscape of AI regulation in the UK is set against the backdrop of national and regional protections for human rights and injects new questions into ongoing debates over whether freedom of expression is adequately protected in the country. These questions are informed by the UK's unique constitutional framework, which lacks a codified, entrenched constitution. This means that the UK constitution is not written down in a single document and that constitutional law does not enjoy a special status in the broader legal framework.[1]

While the right to freedom of expression did not historically enjoy distinct constitutional status, English courts have long recognized that such a right exists in the common law.[2] In the contemporary UK, freedom of expression is protected by way of the Human Rights Act 1998 (HRA), which is part of the country's constitutional framework and incorporates the rights set out in the European Convention on Human Rights (ECHR or Convention) into domestic law. This is because of the UK's status as a member state of the Council of Europe, an international organization with the mission of promoting democracy, human rights, and the role of law across Europe. Article 10 of the ECHR protects the right to freedom of expression. Paragraph 1 of Article 10 lays out the general protection for expression, which includes the freedom to hold opinions and to receive and impart information and ideas without government interference.

The ECHR underpins the Council of Europe's work, the implementation of which is overseen by the European Court of Human Rights (ECtHR). The HRA empowers individuals in the UK to bring cases before domestic courts to vindicate Convention rights and directs domestic courts to, among other things, "take into account" judgments of the ECtHR.[3] One of the key limitations of the HRA in the context of AI regulation is that it only applies to public

1 For a detailed examination of the UK's unique constitutional framework, see Anthony King, *The British Constitution* (OUP, 2007).

2 See, e.g., *Attorney-General v. Guardian Newspapers Ltd* (No. 2) (1990) 1 AC 109, 283-84 (in which Lord Goff opined that there was, in principle, no difference between English law and Article 10 with respect to the right to freedom of expression).

3 HRA 1998, Section 2.

authorities (and those performing a public function).[4] As a result, it does not bestow duties or obligations on private actors. Additionally, the UK ratified the International Covenant on Civil and Political Rights in 1976, which includes the protection of freedom of expression contained in Article 19.

Also relevant to the protection of freedom of expression in the UK is devolution. This is the term used to describe the transfer of power in certain areas from the UK Parliament in Westminster to bodies in Scotland, Wales, and Northern Ireland (the "devolved nations"). The devolved nations are required to comply with the ECHR by virtue of the HRA and specific provisions set out in the devolution statutes.[5]

Regarding the development of AI regulation, at the time of writing, the UK does not have legislation covering AI as a technology, and there is no single AI regulator in the country. This is in contrast to the EU, which has adopted legislation specifically directed to regulating AI that lays out legally binding standards for its use.[6] In the absence of a statutory framework specifically targeting AI in the UK, it is legally regulated according to how it is used.[7] As a result, the use of AI may fall under various statutory frameworks that are overseen by existing regulators covering specific sectors. These regulators include Ofcom (the communications regulator), Ofgem (the energy regulator), and the Financial Conduct Authority (the conduct regulator for financial services). The treatment of AI is also informed through a collection of non-statutory (and non-binding) principles for AI governance, which are examined in Section 3, below.

4 Public authorities include government departments, local authorities, state schools, and police forces.

5 See Section 6(2)(c) and Section 24(1)(a) of the Northern Ireland Act 1998; Section 29(2)(d) and Section 57(2) of the Scotland Act 1998; Section 81(1) and Section 94(6)(c) of the Government of Wales Act 2006. See also Colm O'Cinneide, "Human Rights, Devolution and the Constrained Authority of the Westminster Parliament," *UK Constitutional Law Association* (March 4, 2013), https://ukconstitutionallaw.org/2013/03/04/colm-ocinneide-human-rights-devolution-and-the-constrained-authority-of-the-westminster-parliament/.

6 For an explanation of the distinction between the approaches of the UK and the EU in the area of AI regulation, see Elizabeth Rough, "AI Regulation in the UK," House of Commons Library (June 10, 2026), https://researchbriefings.files.parliament.uk/documents/CBP-10003/CBP-10003.pdf.

7 Rough, 5.

2. GENERAL STANDARDS OF FREEDOM OF EXPRESSION

As outlined in the Introduction, freedom of expression in the UK is protected under Article 10 of the ECHR, by way of the HRA. Paragraph 2 of Article 10 describes the circumstances in which states may interfere with the right to freedom of expression articulated in Paragraph 1 without violating the Convention. It instructs that exercising the right to freedom of expression carries with it duties and responsibilities and that restrictions on expression are permissible so long as they are prescribed by law and necessary in a democratic society in the interests of particular legitimate aims, which include national security, public order, and the protection of the reputation or rights of others. An interference with freedom of expression therefore only violates Article 10 if the government fails to establish the requirements of Paragraph 2, that is, if the government's identified "legitimate aim" does not prevail over the freedom of expression interests of the speaker.

In the seminal case of *Handyside v. United Kingdom,* the ECtHR held that freedom of expression is "one of the essential foundations of [a democratic] society, one of the basic conditions for its progress and for the development of every man."[8] The Court further opined that Article 10 applies not only to information or ideas that are favorably received but also to those that "offend, shock, or disturb."[9] Notwithstanding these principles, the Court cautioned that it "may be necessary in certain democratic societies to sanction or even prevent all forms of expression which spread, incite, promote, or justify hatred based on intolerance provided that any formalities, conditions, restrictions or penalties imposed are proportionate to the legitimate aim pursued."[10] For this reason, interferences with expression by member states in relation to hate speech, terrorist-related speech, and other forms of extreme speech are permitted so long as they satisfy the requirements of Paragraph 2.

8 *Handyside v. United Kingdom* (1976) ECHR 5, para. 49.

9 *Handyside*, para. 49.

10 *Erbakan v. Turkey*, App. No. 59405/00 (ECtHR, October 6, 2006), para. 56 (internal quotations omitted).

Under the HRA, before a bill can become law, a member of Parliament must either make a statement that the provisions of the bill are compatible with Convention rights or that while they cannot make such a statement, the government advises to proceed with the Bill anyway.[11] The reason that Parliament can proceed with a Bill without declaring it compliant with the HRA is because of parliamentary sovereignty. This is a principle of UK constitutional law that regards Parliament as the supreme legal authority in the UK.[12] Section 6 of the HRA instructs that it is unlawful for a public authority to act in a way that is incompatible with a Convention right. Additionally, Section 3 instructs that legislation must be read and applied by courts in a way that is compatible with Convention rights "so far as it is possible to do so."[13]

Regarding the legal standards for access to information in the UK, the Freedom of Information Act 2000 provides public access to information held by public authorities in England, Wales, and Northern Ireland, and by UK-wide public authorities based in Scotland. Information held by Scottish public authorities is covered by Scotland's Freedom of Information (Scotland) Act 2002 (FOISA).[14]

Standards for freedom of expression and access to information apply to the use of AI to the extent that it falls within the scope of existing legislation that, pursuant to the HRA, should comply with Article 10 and other Convention rights. Any future legislation specifically directed to regulating AI will also be subject to these standards.

11 See HRA, Section 19.

12 In practice, this means that Parliament can create or repeal any law and that UK courts lack the authority to strike down legislation.

13 HRA, Section 3.

14 The regulation of freedom of information is a devolved power.

3. AI LEGISLATION LANDSCAPE

At the time of writing, the UK's general approach to the regulation of AI is based on the government's 2023 AI White Paper, which outlined a "pro-innovation approach to AI regulation."[15] This approach eschewed AI-specific legislation in favor of a principles-based approach to be implemented by existing regulators applying five cross-sector non-legally binding principles: safety, transparency, fairness, accountability and contestability.[16] This has led to criticism that the UK's approach suffers from a governance gap due to a lack of consistent monitoring and proactive regulation across sectors.[17]

3.1 The UK's Evolving Approach to AI Regulation

In its 2024 election manifesto, the Labour Party signaled a change in course by pledging to "ensure the safe development and use of AI models by introducing binding regulation on the handful of companies developing the most powerful AI models."[18] In March of 2025, the Labour government expressed an intention to introduce AI-specific legislation, advising that it was currently working on legislative proposals and would launch a consultation in "due course."[19] At the time of writing, the government has neither launched such a consultation nor introduced AI-specific legislation.[20]

While the UK lacks legislation specifically directed to regulating AI, the past few years have seen the introduction of several policy-related efforts to leverage the purported opportunities presented by AI. For example, in 2023, the Department for Science,

15 Office for Artificial Intelligence, "A pro-innovation approach to AI regulation" (August 3, 2023), https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper.

16 Office for Artificial Intelligence, "A pro-innovation approach to AI regulation."

17 See Peter Chai, "The UK's AI governance gap," LSE Blogs (June 2, 2026), https://blogs.lse.ac.uk/politicsandpolicy/the-uks-ai-governance-gap/.

18 "Change: Labour Party Manifesto" (2024), https://labour.org.uk/wp-content/uploads/2024/06/Labour-Party-manifesto-2024.pdf.

19 UK Parliament, "Artificial Intelligence: Regulation" (March 31, 2025), https://questions-statements.parliament.uk/written-questions/detail/2025-03-25/41098.

20 While a private member's bill was introduced in the House of Lords in March 2025 that proposed creating a central AI Authority to coordinate oversight, it lacks government backing and, as a result, is unlikely to become law. Private members' bills are public bills introduced by MPs and Lords who are not government ministers.

Innovation and Technology (DSIT)[21] launched the "Incubator for Artificial Intelligence" with the aim of "bringing together expert technical talent to pioneer transformative applications of AI for the public good."[22] Since the Labour government came into power in 2024, it has focused on initiatives promoting innovation in AI to drive growth and productivity across the UK economy.[23] These include the AI Opportunities Action Plan, published in January 2025, which is premised on the idea that the UK should be "the best state partner to those building frontier AI,"[24] and includes an investment of over £2 billion for implementation.[25] Additionally, the aim of the Data and AI Ethics Framework, introduced in December of 2025, is to provide a set of principles and activities to guide the responsible development, procurement, and use of data and AI in the public sector.[26] The framework advises consideration of the "diverse impacts on human rights and fundamental freedoms" that may result from "irresponsibly designed or misused AI technologies."[27]

With respect to the devolved nations, the Scottish government published its "AI Strategy 2026-2031" in May of 2026, which lays out Scotland's goal of "harnessing AI responsibly, for everyone."[28] In terms of regulation, the Strategy emphasizes that "[p]eople in Scotland require certainty in respect of their rights to privacy, the proper safeguarding of personal data and intellectual property and their wellbeing."[29] It announces support for a UK-level AI regulation that stresses the need for ventures and businesses in Scotland to access markets, investment, and opportunities in Europe.[30]

21 This is a department in the UK Parliament.

22 i.AI Incubator for Artificial Intelligence, "About Us," accessed June 22, 2026, https://ai.gov.uk/about/.

23 Rough, 5.

24 Department for Science, Innovation & Technology, "AI Opportunities Action Plan" (January 13, 2025), https://www.gov.uk/government/publications/ai-opportunities-action-plan/ai-opportunities-action-plan.

25 Department for Science, Innovation & Technology, "AI Opportunities Action Plan: One Year On" (January 29, 2026), https://www.gov.uk/government/publications/ai-opportunities-action-plan-one-year-on/ai-opportunities-action-plan-one-year-on.

26 Government Digital Service, "Data and AI Ethics Framework" (December 18, 2025), https://www.gov.uk/government/publications/data-ethics-framework/data-and-ai-ethics-framework#introduction.

27 Government Digital Service, "Data and AI Ethics Framework."

28 Scottish Government, "Scotland's AI Strategy 2026-2031," https://www.gov.scot/binaries/content/documents/govscot/publications/strategy-plan/2026/03/scotlands-ai-strategy-2026-2031/documents/ai-strategy-scotland/ai-strategy-scotland/govscot%3Adocument/ai-strategy-scotland.pdf.

29 Scottish Government, "Scotland's AI Strategy 2026-2031," 47.

30 Scottish Government, "Scotland's AI Strategy 2026-2031," 48.

3.2 A Case Study of AI Regulation under Existing Legislation: The Online Safety Act (2023)

As highlighted above, because there is currently no legislation expressly directed to AI regulation in the UK, AI is legally regulated to the extent its use is covered by existing legislation. The Online Safety Act (2023) (OSA) provides a useful example of how this works in practice. At just over 280 pages, the OSA reflects the aim of making the UK "the safest place in the world to be online"[31] by introducing a sweeping framework targeting myriad online harms in a single piece of legislation. These harms include cyberbullying, harassment, hate speech, fraud, terrorism, disinformation, and a variety of content—both legal and illegal—deemed harmful to children. It does this by, among other things, placing duties of care on regulated services operating in the UK. Regulated services include social media platforms, photo- or video-sharing services and instant messaging apps, as well as search services.

The duties placed on regulated services include a complex collection of safety measures aimed at reducing the risk of harm to all internet users. These include carrying out risk assessments to determine the risk of harm to individuals posed by illegal content or content that is harmful to children on their services, and preventing or minimizing the risks that users of regulated services encounter such content.[32] The Act applies to services even if the companies providing them are outside the UK, so long as they have links to the UK. This includes a service that has a significant number of UK users, if the UK is a target market, or the UK users can access the service and there is a material risk of significant harm to such users.[33]

As the information regulator, Ofcom is tasked with enforcing the compliance of regulated services under the OSA. It has provided extensive guidance to regulated services about how to implement the duties arising under the OSA, including with respect to AI content.

---

31 See Department of Science, Innovation and Technology, "Britain makes internet safer, as Online Safety Bill finished and ready to become law" (September 19, 2023), https://www.gov.uk/government/news/britain-makes-internet-safer-as-online-safety-bill-finished-and-ready-to-become-law.

32 Ofcom, "Deepfake Defences Mitigating the Harms of Deceptive Deepfakes" (July 23, 2024), https://www.ofcom.org.uk/online-safety/illegal-and-harmful-content/deepfake-defences.

33 Department for Science, Innovation & Technology, "Online Safety Act: explainer" (April 24, 2025), https://www.gov.uk/government/publications/online-safety-act-explainer/online-safety-act-explainer.

For example, Ofcom advises online platforms designated as "user-to-user"[34] services that any AI-generated text, audio, images or videos that are shared by users is regulated in the same way as human-generated content.[35] Thus, deepfake fraud material is regulated no differently from human-generated fraud material, and it is irrelevant whether that content was created on the platform where it is shared or uploaded from elsewhere.[36]

Ofcom further advises that the OSA regulates generative AI tools and content in other ways. For example, generative AI tools that enable the search of more than one website and/or database are "search services" within the meaning of the Act. This includes tools that modify, augment or facilitate the delivery of search results on an existing search engine, or which provide "live" internet results to users on a standalone platform.[37] However, under the OSA as enacted, not all AI chatbots are regulated. As Ofcom explains in relevant guidance, chatbots are not subject to regulation under the OSA if they only allow people to interact with the chatbot itself and no other users and do not search multiple websites or databases when giving responses to users (i.e., are not "search services").[38] Whether AI chatbots fall within one or more of these categories depends on functionality, and even where some AI chatbots are within the Act's scope, this does not necessarily mean that all generated content is regulated. For example, images and videos created by a chatbot without searching the internet are not generally in scope.[39]

To address this gap in coverage of AI-generated content, Parliament amended the OSA by way of secondary legislation in 2026.[40] This includes the Crime and Policing Act 2026 (CPA).[41] The CPA amends the OSA by inserting a section empowering the Secretary of State

34 These are sites or apps that allow their users to interact with each other by sharing images, videos, messages, comments or data with other users of the platform, such as sites that allow users to upload or create their own Generative AI chatbots – "user chatbots" – which are also made available to other users.

35 Ofcom, "Open letter to online service providers operating in the UK about how the UK's Online Safety Act will apply to generative AI and chatbots" (November 8, 2024), https://www.ofcom.org.uk/online-safety/illegal-and-harmful-content/open-letter-to-uk-online-service-providers-regarding-generative-ai-and-chatbots.

36 Ofcom, "Open letter to online service providers operating in the UK about how the UK's Online Safety Act will apply to Generative AI and chatbots."

37 In response to a user query about health information, Ofcom advises that a standalone Generative AI tool that can serve up live results drawn from health advice websites and patient chat forums would make it a "search service" regulated by the Act.

38 Ofcom, "Ofcom update: Investigation into X, and scope of the Online Safety Act" (February 3, 2026), https://www.ofcom.org.uk/online-safety/illegal-and-harmful-content/investigation-into-x-and-scope-of-the-online-safety-act.

39 See Ofcom, "Ofcom update: Investigation into X, and scope of the Online Safety Act."

40 Secondary legislation is law created by ministers (or other bodies) under powers given to them by an Act of Parliament. It is used to fill in the details of Acts, which are primary legislation.

41 The CPA received royal assent on April 29, 2026. Royal assent is when the King formally agrees to make a bill into an Act of Parliament (law).

to amend any provision of the OSA for the purpose of minimizing the risks of harm by illegal AI-generated content and the use of AI services for the commission or facilitation of priority offenses under the Act.[42] This category of offenses includes hate speech and speech-related terrorism offenses. The CPA states that such amendments may place duties on providers of AI services in relation to illegal AI-generated content, the design, operation, or use of AI services relating to illegal AI-generated content, and the use of AI services for the commission or facilitation of priority offenses.[43]

From the beginning, human rights advocates expressed concerns regarding the potential deleterious impact on freedom of speech posed by the OSA's new regulatory framework. While the government asserts that the OSA simply "mak[es] sure what is illegal offline is illegal online,"[44] this fails to acknowledge that policing online speech is of a different nature and scale than the offline context, in which public authorities, not private actors, bear the responsibility of enforcing the law. Of particular importance here is the OSA's sanctions regime, which provides an incentive for regulated services to err on the side of over-removing potentially harmful (and/or illegal) content or not hosting it in the first place. In a press release following the passage of the OSA, the DSIT highlighted the severity of the sanctions regime, stating that companies that fail to comply with the Act "will face significant fines that could reach billions of pounds" and "their bosses may even face prison."[45]

This highlights longstanding concerns over displacing decisions regarding the line between legal and illegal speech from the public authorities to private actors absent safeguards such as judicial oversight and due process.[46] Section 1 of the OSA dictates that regulated services be designed and operated such that "users' rights to freedom of expression... are protected."[47] However, this must be considered in light of the severe sanctions regime

42 Priority offenses reflect the most serious and prevalent illegal content and activity, against which companies must take proactive measures to reduce harm.

43 CPA, Section 248, Section 216A(3)-(4).

44 Department of Science, Innovation and Technology, "Britain makes internet safer, as Online Safety Bill finished and ready to become law."

45 See Department of Science, Innovation and Technology, "Britain makes internet safer, as Online Safety Bill finished and ready to become law."

46 See, e.g., Joelle Fiss and Jacob Mchangama, "The Digital Berlin Wall: How Germany (Accidentally) Created a Prototype for Global Online Censorship," *Justitia* (2019), https://globalfreedom ofexpression.columbia.edu/wp-content/uploads/2019/11/Analyse_The-Digital-Berlin-Wall-How-Germ any-Accidentally-Created-a-Prototype-for-Global-Online-Censorship.pdf; Daphne Keller, "Internet Platforms: Observations on Speech, Danger, and Money," Hoover Institution's Aegis Paper Series No. 1807 (2018), https://ssrn.com/abstract=3262936.

47 OSA, Section 1.

highlighted above. And while the UK government had regard for the right to freedom of expression throughout the legislative process, this is not the same as *grounding* a regulatory framework in human rights.[48] For example, as highlighted by the Electronic Frontier Foundation, regulated services only need a reasonable suspicion that content is illegal before removing it from platforms and removals are driven by platform discretion rather than clear legal rules. This makes it impossible to assess whether each removal is proportionate with respect to interferences with freedom of expression.[49]

Given the weak protections for freedom of expression in the OSA, it is reasonable to question whether a statute targeting AI will include sufficient protections for freedom of expression and other rights implicated in this area. The next section examines specific efforts by the UK government to respect human rights in the context of AI regulation.

3.3 AI Regulation and Human Rights

In the context of AI and human rights, specifically, the UK has signed the Council of Europe's Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (AI Convention) in 2024.[50] The AI Convention adopts a risk-based approach to the design, development, use, and decommissioning of AI systems, which requires consideration of any potential negative consequences of the use of such systems.[51] Among other things, it requires countries to establish an independent oversight mechanism to oversee compliance with the Convention, to raise awareness, and to stimulate an informed public debate on AI and human rights. The AI Convention will not be legally binding in the UK unless and until it is ratified.

The Data and AI Ethics Framework introduced in 2025 advises consideration of the "diverse impacts on human rights and fundamental freedoms" that may result from "irresponsibly designed or misused AI technologies."[52] Also in 2025, the UK Parliament's Joint Committee

48 See Eliza Bechtold, "Regulating Online Harms: An Examination of Recent Developments in the UK and the US through a Free Speech Lens," (2024) 16(2) Journal of Media Law 1.

49 Electronic Frontier Foundation and others, "Briefing for MPs: Petition debate on repealing the Online Safety Act" (December 15, 2025), https://www.openrightsgroup.org/app/uploads/2025/12/Online-Safety-Act-Parliamentary-Petition-Debate-Joint-Briefing.pdf.

50 This is the first international legally binding treaty aimed at ensuring the respect of human rights, the rule of law, and democratic legal standards in the use of AI. At present, no European country has ratified the Convention.

51 Council of Europe, "Council of Europe adopts first international treaty on artificial intelligence" (May 17, 2024), https://www.coe.int/en/web/portal/-/council-of-europe-adopts-first-international-treaty-on-artificial-intelligence.

52 Government Digital Service, "Data and AI Ethics Framework."

on Human Rights (JCHR)[53] launched an inquiry (JCHR Inquiry) to examine the threats and opportunities offered by AI for human rights in the UK, as well as "whether existing legal and regulatory frameworks are sufficient to protect human rights and keep pace with AI development."[54] Questions raised by the JCHR Inquiry include to what extent the UK's existing legal framework provides sufficient protections for human rights in relation to AI and who should be held accountable for breaches of human rights resulting from uses of AI.[55] The deadline for submissions was September 5, 2025.

Written and oral submissions to the JCHR Inquiry provide insight into the challenges of proper oversight of human rights in this context. For example, in written evidence, the Equality and Human Rights Commission (EHRC)[56] noted that in 2012, "the then Government Equalities Office established that £17.1million was sufficient for the Commission to fulfil our statutory functions. Over the following years, our budget reduced from £42.8million in 2011/12 to £17.1million 2019-20 and has remained static since." The EHRC further noted that this amounts to a real-terms cut of over 30 percent and there are no indications that it will receive any budgetary increase in the near future.[57]

This builds on the warning from the EHRC in 2023 that while AI would "bring many benefits, a greater focus on human rights and equality and significantly more funding for regulators (such as the EHRC) was urgently needed to manage the potential of the rapidly advancing technology." Indeed, AI has been a priority for the EHRC since 2022 and it recognizes that, along with potential benefits to society, AI brings a wide range of risks of bias and discrimination as well as risks to the adequate protection of human rights.[58]

53 The JCHR is comprised of Members of Parliament. It examines matters relating to human rights in the UK and scrutinises Bills for combability with human rights obligations.

54 UK Parliament, "New inquiry: Human Rights and the regulation of Artificial Intelligence" (July 25, 2025), https://committees.parliament.uk/committee/93/human-rights-joint-committee/news/208676/new-inquiry-human-rights-and-the-regulation-of-artificial-intelligence/.

55 UK Parliament, "New inquiry: Human Rights and the regulation of Artificial Intelligence."

56 The EHRC is the National Human Rights Institution for England and Wales, and for reserved matters in Scotland (i.e., matters that are not devolved to Scotland) and the regulator of equality legislation. It is an independent statutory body.

57 Equality and Human Rights Commission, "Written Evidence Submitted by the Equality and Human Rights Commission" (RAI0076) (October 2025), https://committees.parliament.uk/writtenevidence/149989/html/.

58 See EHRC, "An update on our approach to regulating artificial intelligence" (April 30, 2024), https://www.equalityhumanrights.com/news/news/update-our-approach-regulating-artificial-intelligence.

4. ELECTION AND POLITICAL CONTENT

The ECtHR provides heightened protection for political speech and recognizes the value of satire and artistic speech to democratic deliberation. However, protections for political speech are tempered by the recognition that not all political speech is equally valid and that some forms of restriction are necessary in democratic societies. Questions over the regulation of political speech become more challenging in and around elections.

4.1 Political Speech and the ECtHR

The ECtHR recognizes political speech as integral to democracy, holding that "there is little scope...for restrictions on political speech or debates on matters of public interest."[59] However, it also holds that not all speech is worthy of inclusion in public debate and that states may and should act as the arbiter of such worthiness. By way of example, member states may restrict speech that does "not contribute to any form of public debate capable of furthering progress in human affairs."[60] In the context of satire, the Court recognizes that while it may qualify as political speech, satire can also function as a "guise" for dangerous speech that does not warrant protection under the Convention.[61] In the context of artistic speech, the Court holds that it enjoys a high level of protection under the Convention.[62] This protection "affords the opportunity to take part in the public exchange of cultural, political and social information and ideas of all kinds."[63] Regarding whether a particular expression qualifies as artistic, the Court usually makes a determination on a

59 *Dichand and others v. Austria*, App. no. 29271/95 (ECtHR, 26 May 2002), para. 38.

60 *Otto-Preminger-Institut v. Austria*, App. no. 13470/87 (ECtHR, 1994), para. 49. Similarly, the Council of Europe expressly rejected "absolute liberalism" in the realm of speech regulation in favor of an approach predicated on the notion that not all ideas are deserving of circulation and that the right to express one's ideas may be outweighed by competing societal interests. See Venice Commission, "Report on the Relationship between Freedom of Expression and Freedom of Religion: the Issue of Regulation and Prosecution of Blasphemy, Religious Insult and Incitement to Religious Hatred" (October 23, 2008), https://www.venice.coe.int/webforms/documents/default.aspx?pdffile=CDL-AD(2008)026-e.

61 See *M'Bala v France*, App. no. 25239/13 (ECtHR, October 20, 2015), para. 40.

62 *Jelševar and Others v. Slovenia* (dec.), 2014, para. 33.

63 *Lindon, Otchakovsky-Laurens and July v. France* [GC], 2007, para. 47.

case-by-case basis.[64]

The approach of the Court in the context of political speech reflects an understanding of freedom of expression under the Convention as a "dialogue and a spirit of compromise, necessarily entailing various concessions on the part of individuals or groups of individuals, which are justified in order to maintain and promote the ideals and values of a democratic society."[65] This requires the subordination of individual interests to those of the group in certain cases to ensure "the fair and proper treatment of people from minorities and avoid any abuse of a dominant position."[66] This explains why the ECtHR does not regard viewpoint-based restrictions on expression as raising free speech concerns and regards restrictions on particular categories of expression, such as hate speech, as justified interferences with Article 10 rights.[67] The regulation of hate speech in the UK is examined in Section 5, below.

4.2 Concerns over False Information in and around Elections

In the context of elections and AI, the UK government is attuned to the challenges arising from the increase in false information in and around elections and the adverse impact this may have on democracy. To address these challenges, the Electoral Commission[68] launched a pilot system in April of 2026 for detecting political deepfakes and addressing the growing threat of AI misinformation in and around elections. The system monitors online content for deepfake audio and video intended to mislead voters about the electoral process or distribute false information regarding a candidate.[69] In introducing the pilot system, the Electoral Commission highlighted that during the 2024 UK general election, over half of voters surveyed by the Commission reported seeing misleading information about parties or candidates, and around 25 percent reported seeing or hearing a deepfake.[70] The pilot was

64 European Court of Human Rights, "Key Theme: Article 10, Artistic expression" (February 28, 2026), https://ks.echr.coe.int/documents/d/echr-ks/artistic-expression.

65 See, e.g., *Kuliś and Różycki v. Poland*, App. no. 27209/03 (ECtHR, January 6, 2010), para. 39.

66 *Gough v. United Kingdom*, App. no. 49327/11 (ECtHR, March 23, 2015).

67 *Gough*, para. 168.

68 The Electoral Commission is the statutory, independent body that oversees elections and regulates political finance in the UK.

69 For example, false videos claiming a candidate has withdrawn from a race. The Electoral Commission, "Electoral Commission launches deepfake detection pilot to counter AI misinformation" (April 15, 2026), https://www.electoralcommission.org.uk/media-centre/electoral-commission-launches-deepfake-detection-pilot-counter-ai-misinformation.

70 The Electoral Commission, "Electoral Commission launches deepfake detection pilot to counter AI misinformation."

live during the May 2026 elections in England, Scotland, and Wales.[71] At the time of writing, no details regarding use of the pilot system were publicly available.

Following the 2026 elections in Scotland, the Electoral Commission called for new legal controls over misinformation from AI chatbots.[72] This resulted from an investigation into the election by the think tank Demos, which found that AI services gave voters misinformation in response to 34 percent of the questions posed.[73] As reported by the Guardian, Demos found that AI tools invented fictitious scandals, misinformed voters regarding the date of the election, and erroneously claimed that voters in Scottish elections needed ID in order to vote.[74] In May of 2026, it was reported that the chief executive of the Electoral Commission was pressing ministers to introduce legislation to make AI companies more accountable around election misinformation.[75]

71 The Electoral Commission, "Electoral Commission launches deepfake detection pilot to counter AI misinformation."

72 Severin Carrell, "ChatGPT and other AI bots made huge errors before Scottish election, study finds," *The Guardian*, May 20, 2026, https://www.theguardian.com/technology/2026/may/20/ai-chatbots-chatgpt-replika-grok-gemini-misinformation-scottish-election-demos.

73 Demos, "Electoral Hallucinations: Safeguarding UK elections in the world of LLMs and AI chatbots" (May 20, 2026), https://demos.co.uk/research/electoral-hallucinations-safeguarding-uk-elections-in-the-world-of-llms-and-ai-chatbots/.

74 Carrell, "ChatGPT and other AI bots made huge errors before Scottish election, study finds."

75 Carrell, "ChatGPT and other AI bots made huge errors before the Scottish election, study finds."

5. HATE SPEECH

The legal framework for hate speech in the UK has developed on a piecemeal basis, resulting in a complex collection of criminal sanctions, civil causes of action, and regulatory codes. The UK's extensive counterterrorism framework also includes speech-related offenses, which criminalize expression based on the viewpoint expressed (i.e., speech that the government characterizes as supporting terrorism). Collectively, these offenses raise concerns regarding the adequate protection of political speech online, which are exacerbated by the onerous duties placed on private companies-and the threat of significant sanctions for failing to comply with such duties-under the OSA.

5.1 The Criminalization of Political Speech Classified as "Hate Speech"

The most significant criminal laws directed to hate speech are the stirring up hatred offenses in the Public Order Act 1986. Unlike other hate crime offenses, in which speech is only relevant to the issue of aggravation or sentence enhancement, expression forms the basis of the stirring up hatred offenses. These offenses cover speech involving racial hatred, religious hatred, and hatred on the grounds of sexual orientation. While the stirring up offenses do not explicitly target online speech, courts have recognized that the offenses (along with other criminal offenses that target hateful expression)[76] apply to online content.[77] While there is no freedom of expression exception to the offense of stirring up racial hatred, exceptions exist for hatred based on religion[78] and sexual orientation.[79] The racial hatred offense does not require intent to stir up racial hatred and it is sufficient that

76 See, for example, the new criminal False Communications Offense (Section 179) of the OSA.

77 See *R v. Sheppard* [2010] EWCA Crim 65.

78 See Section 29J, which states that "[n]othing in this Part shall be read or given effect in a way which prohibits or restricts discussion, criticism or expressions of antipathy, dislike, ridicule, insult or abuse of particular religions or the beliefs or practices of their adherents, or of any other belief system or the beliefs or practices of its adherents, or proselytising or urging adherents of a different religion or belief system to cease practising their religion or belief system."

79 See Section 29JA, which states that "[f]or the avoidance of doubt, the discussion or criticism of sexual conduct or practices or the urging of persons to refrain from or modify such conduct or practices shall not be taken of itself to be threatening or intended to stir up hatred."

"racial hatred is likely to be stirred up" by the expression at issue.[80] The offenses for religion and sexual orientation require the intent to stir up hatred, but there is no requirement for the authorities to demonstrate that any hatred was actually, or likely to be, stirred up as a result of the expression.

Scotland elected to legislate in this area by way of the Hate Crime and Public Order (Scotland) Act 2021 (2021 Act). It repeals certain provisions of the Public Order Act 1986 applicable to Scotland and introduces new stirring up hatred offenses based on the following characteristics: race, age, disability, religion (in the case of a social or cultural group, "perceived religious affiliation" qualifies), transgender identity, and variations in sex characteristics.[81] Section 9 of Part 3 lays out the protection for freedom of expression and is notable for its brevity. It provides that behavior or material is not to be taken to be threatening or abusive solely on the basis that it "involves or includes discussion or criticism of matters relating to" all protected characteristics with the exception of race, as well as "discussion or criticism relating to, or expressions of antipathy, dislike, ridicule or insult towards religion," proselytizing, or urging of persons to cease practicing their religions.[82]

5.2 Speech Offenses Characterized as Supporting or Encouraging Terrorism

Part of the UK's extensive counterterrorism legislative framework includes criminal offenses for expression characterized as supporting or encouraging terrorism. This includes an "encouragement of terrorism" offense, which applies in circumstances in which a "reasonable person would understand [a] statement as an encouragement or inducement to them to the commission, preparation or instigation of acts of terrorism."[83] Accordingly, in targeting individuals for prosecution, there is no onus on the government to establish that a reasonable person did, in fact, understand particular expression as such or that any likelihood of harm resulted therefrom. The framework also includes the offense of "Proscription," which criminalizes the expression of an opinion or belief that is supportive

80 See POA, Part 3, Section 18.

81 2021 Act, Part 3.

82 2021 Act, Part 3(9)(b).

83 Terrorism Act 2006, Section 1.

of a proscribed organization or inviting support for a proscribed organization.[84] Legislation empowers the Home Secretary to proscribe an organization if "they believe it is concerned in terrorism, and it is proportionate to do."[85]

5.3 The ECtHR's Permissive Approach to Criminal Laws Targeting Political Speech

In its approach to adjudicating hate speech cases, the ECtHR has demonstrated a tolerance for interferences based on the legitimate aim of protecting the rights and reputation of others, that is, the groups targeted by this type of expression. The Court has elected not to adopt the Council of Europe's definition of "hate speech,"[86] opting instead for an "autonomous conception" of hate speech and analyzing alleged interferences on a case-by-case basis.[87] In cases involving an Article 10 challenge to interference with purported hate speech, the Court analyzes whether the expression at issue falls into one of two categories. The first category is "comprised of the gravest forms of hate speech," which the Court construes as falling under Article 17 and "thus entirely excluded from the protection of Article 10."[88] Other types of extreme speech that fall into this category, according to the Court, are antisemitism and Holocaust denial,[89] attacks against religious groups,[90] and advocating for National Socialism.[91]

The second category is "comprised of 'less grave' forms of hate speech" that the Court does not regard as falling entirely outside the protection of Article 10, but which it has considered

84 UK Home Office, "Proscribed terrorist groups or organisations" (July 11, 2025), https://www.gov.uk/ government/publications/proscribed-terror-groups-or-organisations--2/proscribed-terrorist-groups-or- organisations-accessible-version.

85 See UK Home Office, "Proscribed terrorist groups or organisations."

86 The Council of Europe defines "hate speech" as constituting "all forms of expression which spread, incite, promote or justify racial hatred, xenophobia, antisemitism or other forms of hatred based on intolerance, including intolerance expressed by aggressive nationalism or ethnocentrism, discrimination and hostility against minorities, migrants, and people of immigrant origin." Recommendations and Declarations of the Committee of Ministers of the Council of Europe in the field of media and information society, *Directorate General of Human Rights and Rule of Law* (2016), 76-78, outlining Recommendation No R (97) 20 of the Committee of Ministers to member states on "hate speech" (adopted 30, October 1997).

87 Francoise Tulkens, "When to say is to do: freedom of expression and hate speech in the case-law of the European Court of Human Rights," ECtHR European Judicial Training Network: Seminar on Human Rights for European Judicial Trainers (October 9, 2012), 2.

88 *Lilliendahl v. Iceland*, App. no. 29297/18 (ECtHR, June 11, 2020), para. 5. Article 17 is viewed as the Convention's "anti-abuse clause." It provides that the ECHR does not imply a right for any state, group or person "to engage in any activity or perform any act aimed at the destruction of any of the rights and freedoms set forth herein or at their limitation to a greater extent than is provided for in the Convention."

89 See *M'Bala v. France*, App. no. 25239/13 (ECtHR, October 20, 2015).

90 See *Norwood v. United Kingdom*, App. no. 23131/03 (ECtHR, November 16, 2004).

91 See *Kühnen v. Federal Republic of Germany*, App. no. 12194/86 (ECtHR, May 12, 1988).

permissible for member states to restrict.[92] Included in this category is not only speech that explicitly calls for violence or other criminal acts, but also "attacks on persons committed by insulting, holding up to ridicule or slandering specific groups of the population," which "can be sufficient for allowing the authorities to favour combating prejudicial speech within the context of permitted restrictions on freedom of expression."[93] In cases concerning speech that does not call for violence or other criminal acts, but which the Court considers to constitute hate speech, its conclusion is based on an assessment of the content of the expression and the manner in which it is delivered.[94]

5.4 Concerns Regarding the Adequate Protection of Political Speech in the UK

NGOs, the UN, and legal scholars have expressed concern regarding the breadth of the UK's statutory framework of speech-related offenses and the likelihood of chilling political speech that does not incite violence or hatred.[95] A recent example of the negative impact on freedom of expression resulting from this framework is the crackdown on protests in support of Palestinian human rights in the context of the war in Gaza. In 2025, the government added Palestine Action to the list of proscribed organizations under the Terrorism Act 2000, notwithstanding the urging of United Nations (UN) experts not to do so.[96] This means that anyone expressing support for Palestine Action is now subject to criminal prosecution under the UK's Proscription offense. UN experts warned that designating Palestine Action, which "positions itself as a national activist network that promotes civil disobedience and takes direct action against companies and institutions deemed by the group as being used by Israel to violently enforce apartheid, occupation, colonisation and genocide in Palestine" would criminalize legitimate activities by innocent members of the group.[97] The UK ignored these warnings and police arrested more than a

92 *Lilliendahl v. Iceland*, App. no. 29297/18 (ECtHR, June 11, 2020), para. 33.

93 *Lilliendahl*, para. 36.

94 *Lilliendahl*, para. 36.

95 See, e.g., Amnesty International, "United Kingdom: End Prosecution of Peaceful Protesters" (November 22, 2025), https://www.amnesty-usa.org/urgent-actions/united-kingdom-end-prosecution-of-peaceful-protestors/; UN News, "New UK law curtails key civil and political rights: UN rights chief" (April 27, 2023), https://news.un.org/en/story/2023/04/1136107.

96 Office of the United Nations High Commissioner for Human Rights, "UN experts urge United Kingdom not to misuse terrorism laws against protest group Palestine Action" (July 1, 2025), https://www.ohchr.org/en/press- releases/2025/07/un-experts-urge-united-kingdom-not-misuse-terrorism-laws-against-protest.

97 Office of the United Nations High Commissioner for Human Rights, "UN experts urge United Kingdom not to misuse terrorism laws against protest group Palestine Action."

thousand people in the summer of 2025, many of whom were peaceful protesters in their sixties and seventies, for violating the country's counterterrorism laws.[98]

Similarly, the ECtHR has garnered criticism for its failure to adopt a definition of "hate speech" and its unwillingness to apply a precise test for adjudicating Article 10 cases involving hateful expression, which has resulted in fragmented and methodologically underdeveloped doctrine that fails to adequately protect the targets of national efforts to criminalize hateful expression.[99] Given the UK's extensive framework for criminalizing political speech, the ECtHR's lack of rigor is particularly concerning.

All of these concerns are exacerbated in the context of AI, as legislative frameworks like the OSA prioritize the targeting of purported hate speech and terrorist-related expression absent the due process and other protections afforded in the offline context. Moreover, sanctions regimes may increase reliance on unreliable AI systems to monitor and remove content, threatening the type of political speech that is essential to healthy democracies.[100]

98 See Caroline Davies, "'At 80, to be treated like a terrorist is shocking': arrested on suspicion of supporting Palestine Action," (July 30, 2025), *The Guardian*, https://www.theguardian.com/politics/2025/jul/30/at-80-to-be-treated-like-a-terrorist-is-shocking- arrested-on-suspicion-of-supporting-palestine-action.

99 For an in-depth and critical analysis of the ECtHR's hate speech jurisprudence, see Natalie Alkiviadou and Jacob Mchangama, "Hate Speech and the European Court of Human Rights: Whatever Happened to the Right to Offend, Shock, or Disturb?," (2021) 00 H.R.L Rev. 1. In this piece, the authors argue that the Court has adopted an overly-restrictive approach to hate speech that fails to adequately protect controversial political speech while also failing to adopt an adequate test for balancing the impact of hate speech with freedom of expression and simply assuming, rather requiring member states to demonstrate, that prohibitions on hate speech are the most efficient tool to foster tolerance and social cohesion.

100 See European Agency for Fundamental Rights, "Bias in Algorithms: Artificial Intelligence and Discrimination" (2022), https://fra.europa.eu/sites/default/files/fra_uploads/fra-2022-bias-in-algorithms_en.pdf.

6. DEFAMATION

Devolution is particularly relevant to the regulation of defamation in the UK. In England and Wales, defamation is regulated under the Defamation Act 2013 (Defamation Act). The purpose of the Defamation Act, according to the government, was to "to rebalance the law on defamation to provide more effective protection for freedom of speech while at the same time ensuring that people who have been defamed are able to protect their reputation."[101] Pursuant to this aim, Section 5 of the Defamation Act introduced a new defense for website operators hosting user-generated content. This defense applies if a website operator can show that it did not post the statement on the website. However, the defense is not applicable if the person bringing the claim can show that it was not possible for the claimant to identify the person who posted the statement, that they gave the website operator a notice of complaint in relation to the statement, and that the operator failed to respond to that notice in accordance with the procedure set out in the Defamation (Operators of Websites) Regulations 2013.[102]

In Scotland, defamation is regulated under the Defamation and Malicious Publication (Scotland) Act 2021, which, like the Defamation Act 2013, includes a defense for website operators.[103] The Northern Ireland Executive refused to extend the 2013 Act to Northern Ireland, thus distinguishing Northern Ireland from England and Wales in this context.[104] Accordingly, Northern Ireland's legislation, the Defamation Act (NI) 2022, differs from the Defamation Act 2013 by, among other things, omitting regulations relating to website operators. This has prompted criticism from commentators that the Act does not adequately address the continuing challenges presented by online publication.[105]

---

101 Ministry of Justice, "Complaints about defamatory material posed on website: Guidance on Section 5 of the Defamation Act 2013 and Regulations" (January 2014), https://assets.publishing.service.gov.uk/media/5a7c606440f0b626628aba0b/defamation-guidance.pdf.

102 The Defamation (Operators of Websites) Regulations 2013. This involves drafting a Notice of Complaint.

103 See Section 5.

104 See Alyssa Andrade Palmieri, "The Deficiency in Defamation Laws in Northern Ireland and the Protection of Human Rights," Queens University Belfast (March 30, 2022), https://blogs.qub.ac.uk/studentlawjournal/2022/03/30/1106/.

105 See, e.g., Department of Finance, "Review of the Defamation Act (NI) 2022" (June 2024), https://www.finance-ni.gov.uk/sites/default/files/publications/dfp/Review%20of%20the%20Defamation%20Act%20%28NI%29%202022%20published.pdf.

7. PROTECTION OF MINORS

The UK is making a concerted effort to strengthen protections for children online through both legislation and policy initiatives. For example, the OSA places expansive obligations on regulated services to assess the risk posed by content—both legal and illegal—that is harmful to children. Ofcom advises that this includes some forms of deepfake technology, including deepfake child sexual abuse material (CSAM). Regulated services also have a duty to take steps to prevent or minimize the risk of their users encountering this kind of content. Additionally, the CPA introduces a new offense that criminalizes AI models that have been optimized to create child sexual abuse material CSAM, which produce hyperrealistic CSAM that often contains the likeness of real children.[106]

The OSA also requires a regulated service to use age verification[107] or age estimation[108] (or both) to prevent children of any age from encountering "primary priority content that is harmful to children" that it identifies on the service.[109] This category includes pornographic content and content that encourages suicide or an eating disorder. As highlighted above in Section 3, the OSA covers generative AI tools that can generate pornographic material, and services are required to use "highly effective age assurance" to ensure children cannot normally access this type of content.[110]

In the policy realm, the Labour government ran a consultation titled "Growing up in the online world" from March to May of 2026. Noting that the OSA established "one of the strongest systems in the world for protecting children from illegal and harmful online content," the government consultation considered the broader impact of technology on children's everyday lives and whether additional measures could build on the OSA to ensure

106 Ministry of Justice, "Crime and Policing Act 2026: child sexual abuse material factsheet" (May 19, 2026), https://www.gov.uk/government/publications/crime-and-policing-act-2026-factsheets/crime-and-policing-act-2026-child-sexual-abuse-material-factsheet.

107 Defined as "any measure designed to verify the exact age of users of a regulated service." OSA Section 230(2).

108 Defined as "any measure designed to estimate the age or age-range of users of a regulated service." OSA Section 230(3).

109 OSA, Section 61(1)-(5).

110 Ofcom, "Open letter to online service providers operating in the UK about how the UK's Online Safety Act will apply to Generative AI and chatbots."

children's safety online.[111] Among other things, the consultation considered the imposition of age restrictions on social media and other services, including AI chatbots.

In June of 2026, the Secretary of State published a report on the consultation for Parliament pursuant to the Children's Wellbeing and Schools Act 2026 (CWSA).[112] The CWSA amended the OSA to provide the Secretary of State with the authority to introduce new legally binding duties on regulated services to prevent or restrict children's access to specific internet services. The report laid out the government's plan to ban social media companies from providing their services to children under the age of 16 and to prevent those under 18 from accessing AI chatbot services that primarily offer sexualized content. That same month, the Labour government announced its aim of passing regulation "before Christmas" to ban children under the age of 16 from accessing a range of social media platforms.[113] At the time of writing, it is unclear whether the UK's new prime minister, Andy Burnham, will move forward with the government's plans in this context.

111 Department for Science, Innovation & Technology, "Growing up in the online world: a national conversation" (June 15, 2026), https://www.gov.uk/government/consultations/growing-up-in-the-online-world-a-national-consultation/growing-up-in-the-online-world-a-national-conversation.

112 Department for Science, Innovation & Technology, "Growing up in the online world: progress statement" (June 2026), https://assets.publishing.service.gov.uk/media/6a2fbb99d95ffddb05d4b014/UAP1_2205_PR.pdf.

113 At the time of writing, the government has not announced the full list of platforms that will be subject to the ban. See Imran Rahman-Jones et. al., "Under-16s will be banned from social media from early 2027," *BBC*, June 15, 2026, https://www.bbc.co.uk/news/articles/c2kydl0zqeko.

8. EXPLICIT CONTENT

The regulatory landscape in the UK extends to explicit content online, with the most recent addition to statutory restrictions introduced by way of the Data (Use and Access) Act 2025 (DUAA). The DUAA introduced new criminal offenses that target the creation of deepfakes. Specifically, it amends the Sexual Offences Act 2003 to add a criminal offense for intentionally creating explicit deepfakes of an adult without their consent. It further criminalizes requesting the creation of these types of deepfakes.[114]

114 See Department for Science, Innovation & Technology, "Data (Use and Access) Act 2025: Explanatory Notes" (2025), https://www.legislation.gov.uk/ukpga/2025/18/pdfs/ukpgaen_20250018_en.pdf.

9. COPYRIGHT

Under the Copyright, Designs and Patents Act 1988 (CDPA), copyright protection is available for any literary, dramatic, musical, or artistic work that is "generated by a computer in circumstances such that there is no human author."[115] In terms of copyright ownership for computer-generated works, the "author" of such a work is deemed to be the person "by whom the arrangements necessary for the creation of the work are undertaken."[116] In the case of a general-purpose AI that generates output in response to a user prompt, the "author" will usually be the person who entered the prompt.[117]

Protected works under the CDPA include books, news articles, computer programs, photographs, musical works, films, and broadcasts. Copyright arises automatically in qualifying works, as defined under the Act, and gives rightsholders exclusive control over a number of "restricted" acts, including copying. There is currently no broad exception in UK law that explicitly permits the use of copyrighted works for commercial text and data mining or the training of generative AI models.

The AI Opportunities Action Plan makes several recommendations intended to encourage the expansion of the UK's AI sector, including that the government should seek to increase the amount of data available to AI developers. On the issue of regulating copyright and AI more generally, the government highlighted the aim of ensuring the UK has the "right regulatory regime that addresses risks and actively supports innovation will drive AI trust and adoption across the economy."[118]

115 Copyright, Designs and Patents Act 1988 (CDPA), Section 178.

116 CDPA, Section 9(3). See also Intellectual Property Office, "Report on Copyright and Artificial Intelligence" (March 18, 2026), https://www.gov.uk/government/publications/report-and-impact-assessment-on-copyright-and-artificial-intelligence/report-on-copyright-and-artificial-intelligence#fn:63.

117 Intellectual Property Office, "Report on Copyright and Artificial Intelligence."

118 Edward Scott, "Copyright and artificial intelligence: Impact on creative industries," *House of Lords Library* (January 27, 2025), https://lordslibrary.parliament.uk/copyright-and-artificial-intelligence-impact-on-creative-industries/#heading-4.

As part of its ongoing work concerning copyright and AI, the Labour government ran a consultation on possible reform of UK copyright law from December 2024 to February 2025. The consultation characterized the country's position on AI and copyright as one of "legal uncertainty" for AI developers, noting that the growing use of realistic impersonation through AI creates new risks—as well as opportunities—for artists and the general public.[119] Following the results of this consultation, the Intellectual Property Office published a 137-page report considering the use of copyright works in the development of AI.[120] No firm proposals flowed from this report. Instead, it highlights that, while there is no prevailing view as to what reform in this area should look like, the responses to the consultation demonstrated growing concern from the creative industries, especially performers, and a clear concern that the existing legal framework may be insufficient to adequately protect a person's image and voice.[121] For this reason, the government expressed interest in considering whether a new personality right may be appropriate.[122]

119 Intellectual Property Office, "Report on Copyright and Artificial Intelligence."

120 The consultation ran from December to 2024 to February 2025.

121 Intellectual Property Office, "Report on Copyright and Artificial Intelligence."

122 Intellectual Property Office, "Report on Copyright and Artificial Intelligence."

10. AI-POWERED SURVEILLANCE

The expanding use of AI in the context of surveillance is a growing concern among human rights advocates across the UK. According to the Oxford Institute of Technology and Justice (Oxford Institute), "[a]t every stage of the criminal process in the UK, at least one actor is using AI in some form."[123] This includes the use of AI for predictive analytics and for facial recognition. Despite this widespread use of AI, there are no statutory regulations directed at its use by law enforcement or others working in the criminal justice system. In the policy realm, the UK government published a white paper titled "From local to national: a new model for policing" in January of 2026, which provides a policy blueprint for policing reform across the UK.[124] This blueprint includes the expansion of facial recognition technologies and the creation of a National Centre for AI and Policing.[125]

In June of 2026, the government announced the launch of "PoliceAI," which it describes as "a new national centre dedicated to the responsible development, piloting and scaling of artificial intelligence in policing."[126] The center will work across all police forces in England and Wales to identify, test, and scale AI tools.[127] It is part of a £140 million investment over 3 years, which includes funding for, among other things, forty live facial recognition units.[128]

123 Oxford Institute of Technology and Justice, "United Kingdom" (March 2026), https://www.techandjustice.bsg.ox.ac.uk/research/united-kingdom.

124 Home Office, "From local to national: a new model for policing" (January 2026), https://www.gov.uk/government/publications/from-local-to-national-a-new-model-for-policing/from-local-to-national-a-new-model-for-policing-accessible. This model applies to England and Wales.

125 Home Office, "From local to national: a new model for policing."

126 Home Office, "From local to national: a new model for policing."

127 Home Office, "PoliceAI to speed up investigations and fight crime" (June 10, 2026), https://www.gov.uk/government/news/policeai-to-speed-up-investigations-and-fight-crime.

128 Home Office, "PoliceAI to speed up investigations and fight crime."

11. MEASURES EMPOWERING FREEDOM OF EXPRESSION

Measures aimed at empowering individuals to use AI in ways that promote engagement with, and participation in, public discourse are being introduced by the UK government as well as the governments of the devolved nations. For example, at the UK level, the DSIT is leading the Media Literacy Action Plan (Action Plan) on behalf of the government. The Action Plan sets out the steps that government departments are taking to strengthen media literacy from 2026 to 2029, which include helping people across the UK "think critically about online content, find trustworthy information, and make the most of digital opportunities."[129] It also lays out specific priorities, including boosting local initiatives to support people facing barriers to participation in online spaces.[130]

In the context of the devolved nations, one of the primary aims of the Welsh government's "Cymraeg [Welsh] 2050 strategy" is "to ensure that the Welsh language is at the heart of innovation in digital technology."[131] To achieve this aim, the Welsh government is committing to "transforming the Welsh-language digital landscape."[132] This strategy includes the specific goal of improving Welsh-language AI and speech-language technologies.[133] The Welsh government highlights that an important part of its Welsh language technology work is helping people with specific needs, including those with disabilities and with additional learning needs.[134] This includes, for example, offering Welsh and bilingual voices for screen

129 Department for Science, Innovation & Technology, "A Safe, Informed Digital Nation: government's Media Literacy Action Plan (2026-2029)" (March 2026), https://www.gov.uk/government/publications/a-safe-informed-digital-nation, 5.

130 Department for Science, Innovation & Technology, "A Safe, Informed Digital Nation: government's Media Literacy Action Plan (2026-2029)," 9.

131 Welsh Language Commissioner, "Artificial Intelligence and the Welsh language: The Welsh Language Commissioner's Regulatory Policy Statement" (2026), https://www.welshlanguagecommissioner.wales/media/hscpv33c/20250722-datganiad-polisi-rheoleiddiol-deallusr-wydd-artiffisial-ar-gymraeg-saesneg.pdf, 1.

132 Welsh Language Commissioner, "Artificial Intelligence and the Welsh language: The Welsh Language Commissioner's Regulatory Policy Statement," 1.

133 Welsh Government, "Welsh language technology and AI" (last updated November 18, 2025), https://www.gov.wales/welsh-language-technology-and-ai.

134 Welsh Government, "Welsh language technology and AI."

readers that blind and visually impaired people use.[135]

The Scottish government's efforts include participating in a project exploring children's rights and AI, which examined real-world applications of AI in primary schools across Scotland and how they impacted children's human rights.[136] Additional initiatives include a free AI course that is designed to help people to understand the impact of AI on daily life, highlights ethical considerations and encourages responsible use. The course is intended to "empower[] educators, pupils, and the wider public to engage confidently with AI technologies at their own pace."[137]

Finally, the Northern Ireland government is also investigating how AI can be leveraged to support children. For example, it is conducting a research study to evaluate how AI tools can support literacy among disadvantaged primary school students as well as those with special educational needs as part of its RAISE program.[138] The program is described as "a whole community and place-based approach in localities across Northern Ireland to support the Department of Education's vision that 'every child is happy, learning and succeeding.'"[139] Its strategic objectives include raising achievement and addressing educational disadvantages.

135 Welsh Government, "Welsh language technology and AI: updates" (last updated September 3, 2025), https://www.gov.wales/welsh-language-technology-and-ai-updates.

136 Scottish AI Alliance, "Trustworthy, ethical and inclusive (2021-2026)," https://static1.squarespace.com/static/5dc00e9e32cd095744be7634/t/69bd568cb10020129d46d31b/1774016140787/SAIA-5-Year-Report_Spreads_ONLINE.pdf.

137 Scottish Government, "Artificial Intelligence (AI) in schools: guidelines and guardrails" (March 25, 2026), https://www.gov.scot/binaries/content/documents/govscot/publications/advice-and-guidance/2026/03/scottish-guidelines-guardrails-use-artificial-intelligence-ai-schools/documents/scottish-guidelines-guardrails-use-ai-schools/scottish-guidelines-guardrails-use-ai-schools/govscot%3Adocument/scottish-guidelines-guardrails-use-ai-schools.pdf. See also https://livingwithai.me/.

138 Department of Education, "Education Minister announces research study evaluating how AI tools can support literacy" (June 2, 2025), https://www.education-ni.gov.uk/news/education-minister-announces-research-study-evaluating-how-ai-tools-can-support-literacy.

139 Department of Education, "RAISE Programme – summary document" (May 31, 2024), https://www.education-ni.gov.uk/publications/raise-programme-summary-document.

12. CONCLUSION

Successive governments in the UK, led by the Conservative and Labour parties, have demonstrated a strong commitment to leveraging the opportunities created by AI to benefit various sectors of the government and UK society more broadly. This includes promised legislation specifically directed to regulating AI, significant financial investment in AI as a key component of broad-based economic growth, and a suite of cross-sector policy initiatives.

Rather than adequately funding the regulators in charge of ensuring policy and legislative developments comply with human rights standards, the government has elected to invest billions of pounds to further its aim of harnessing the benefits of AI.[140] And while the JCHR Inquiry is a positive sign that the government is aware of the potential threats to human rights, including freedom of expression, arising from the development and use of AI, the overall emphasis of policy initiatives and investment is on the potential economic advantages to be gained from increased use and reliance on AI throughout the private and public sectors. Moreover, if the government's commitment to protecting human rights while developing a broader AI strategy can be measured by its investment in the EHRC, the independent statutory body responsible for protecting the human rights of those living in the UK, it is reasonable to conclude that the UK is not taking a human rights-based approach to regulation in this area and that freedom of expression is unlikely to be adequately protected as the government moves forward with its ambitious plans to "shape the AI revolution."[141]

140 Department for Science, Innovation & Technology, "AI Opportunities Action Plan: One Year On."

141 Department for Science, Innovation & Technology, "AI Opportunities Action Plan."



THE FUTURE OF FREE SPEECH

www.futurefreespeech.org