DEMOS

A DECLARATION ON DIGITAL RIGHTS

EMBEDDING HUMAN RIGHTS IN A NEW DEAL FOR THE DIGITAL AGE

SOFIA LYALL
CARAGH AYLETT-BULLOCK
NAEMA MALIK
NICOLA STOKES
ABDURAMAN SESAY
TYREESE CALNAN
ELIZABETH SEGER

FEBRUARY 2026

Published by Demos February 2026

15 Whitehall, London, SW1A 2DD
T: 020 3878 3955
hello@demos.co.uk
www.demos.co.uk

CONTENTS

ACKNOWLEDGEMENTS

We would like to thank all those who contributed to our research, including representatives from government, regulators, digital rights organisations and academics who participated in our interviews and workshop discussions. We would particularly like to thank the following individuals and organisational representatives who generously shared their time and expertise:

- Ada Lovelace Institute
- Amnesty International (Amnesty Tech)
- Article19
- Dr Joe Atkinson, Uni. of Southampton
- AWO
- James Ball, Demos Fellow
- Big Brother Watch
- Dr Elinor Carmi, City University
- Connected by Data
- Defend Digital Me
- Digital Freedom Fund
- Digital Poverty Alliance
- Electronic Freedom Fund
- Equality Now/AUDRi
- European Centre for Not-for-Profit Law
- Professor Ann Kristin Glenster
- Good Things Foundation
- Institute for the Future of Work
- ICRAC
- Sarah Kiden, Responsible AI UK
- Liberty
- Dr Kerry McInerney, Leverhulme Centre for the Future of Intelligence
- Migrants' Rights Network
- Minderoo Centre for Technology & Democracy
- Dr Daragh Murray, Queen Mary, University of London
- NSPCC
- Open Rights Group
- Oxford Martin School of AI Governance
- Dr Yulu Pi, Research Center Trustworthy Data Science and Security
- Public Law Project
- Royal Society
- Dr Birgit Schippers, Uni. of Strathclyde
- Professor Elke Schwarz, Queen Mary, University of London
- Statewatch
- Trades Union Congress
- WITNESS
- Worker Info Exchange
- 5rights

This report was produced by Demos and is editorially independent. Legal and policy analysis was conducted by Sumaya Nur Adan from the Oxford Martin AI Governance Initiative. Any errors are the authors' own.

We would like to thank the generous support of the National Information Society Agency (NIA) of the Government of the Republic of Korea who funded this work between May 2025 and January 2026.

Thank you also to our colleagues for their input and support at various stages

throughout this project: Hannah Perry, Jamie Hancock, Polly Curtis, Sumaya Akthar, Lottie Skeggs, and Chloe Burke.

Sofia Lyall

Caragh Aylett-Bullock

Naema Malik

Nicola Stokes

Abduraman Sesay

Tyreese Calnan

Sumaya Nur Adan

Elizabeth Seger

February 2026

ABOUT THIS REPORT

Demos is Britain's leading cross-party think tank. We put people at the heart of policy-making to create bold ideas and a more collaborative democracy. Our vision is an upgraded democracy, powered by trusting relationships, information and technology, fit for our times. To upgrade democracy, we need a new deal to repair the broken systems that are undermining democracy. The new deal asks something of us all, delivers for all citizens and is built through collaboration. This paper is part of Demos' strategic focus area on '***Trustworthy Technology***' which sees technology and AI as offering make or break opportunities for the new deal: a new deal not just between citizens and the state, but also between citizens and private institutions, namely technology companies.

In this report, we make the case that a demonstrated commitment from the government to preserving human rights in the face of rapid technological change will be critical for a new deal in the digital age. As such, we highlight how data-driven technologies are currently undermining seven key human rights and provide recommendations for the government on how to rebuild protections for these rights in order to rebuild trust between the citizen and the state in the digital age. Key among the recommendations is that the UK Government commit to a 'Declaration of Digital Rights' alongside binding and enforceable human rights-based tech regulation applicable to both public and private actors. The report develops on our 2025 report titled '*Advancing Digital Rights in 2025: Trends, Challenges, and Opportunities in the UK, EU, and Global Landscape*.'[1]

This report was funded by the National Information Society Agency (NIA), of the Government of the Republic of Korea following the publication of their Digital Bill of Rights in 2023. Our work remains editorially independent.

1 Perry, H. et al. (2025). Advancing Digital Rights in 2025: Trends, Challenges, and Opportunities in the UK, EU, and Global Landscape. Demos. https://demos.co.uk/wp-content/uploads/2025/02/Digital-Rights-in-2025.ac_.pdf

EXECUTIVE SUMMARY

Digital technologies, algorithms, and artificial intelligence now shape almost every aspect of social, economic, and political life in the UK. From access to public services and employment to policing, migration control, and democratic participation, decisions that affect people's lives are increasingly mediated by data-driven systems. While these technologies offer opportunities for innovation, efficiency, and inclusion, there is mounting evidence that they also pose significant risks to fundamental human rights. These risks are often felt most acutely by already marginalised communities, embedding and amplifying existing inequalities in a digitally mediated society.

In the face of rapid technological change, we urge the UK government to make a firm commitment to people's human rights in the digital age. At Demos we believe a new deal is needed between citizens and state to rebuild a functional collaboration in which the state listens, citizens trust, and government is able to deliver more effective policy with resulting public support. A commitment to preserving human rights will be critical to this mission, demonstrating to people that their interests are at the heart of government policy priorities and facilitating smoother and positive tech transitions as a result. Specifically, we recommend the government make a bold move in supporting the development of a Declaration of Digital Rights, and committing to it. We provide an example Draft Declaration in the appendix.

This report begins with an assessment of the impacts of AI and data-driven digital technologies on human rights in the UK at a critical moment for technology governance.

The introduction starts by situating the UK within a global context in which rights-based approaches to digital regulation (such as the EU AI Act, the UN Global Digital Compact, and the Council of Europe's Framework Convention on AI) are increasingly being challenged by a growing deregulatory agenda. The deregulatory agenda is being driven by geopolitical competition, security priorities, and corporate influence. In the UK, the absence of a comprehensive cross-sectoral framework for AI, combined with the weakening of existing data protection safeguards, risks leaving significant gaps in human rights protection and further undermining public trust in government tech policy decisions and initiatives.

Part 1 of the report then brings together insights from across the UK digital rights community to examine how AI, algorithms, and data-driven technologies impact seven key human rights drawn from international human rights frameworks. The chapters focus on:

1. **the right to equality and non-discrimination:** AI and algorithms can reinforce existing societal discrimination against marginalised communities because training datasets may embed social biases. Additionally, content moderation and recommender systems on online platforms may fail to effectively moderate, or amplify, discriminatory content.
2. **the right to privacy:** Some digital technology use cases by state and private actors can challenge privacy and data protection rights because they rely on indiscriminate, unnecessary, and illegitimate data processing. This holds implications for other rights, including non-discrimination, and freedom of expression.
3. **the right to freedom of expression and information:** Ineffective systems and processes on online platforms, which effectively function as online public spaces may restrict the right to access information. Surveillance technologies such as facial recognition or crime prediction may lead to chilling effects on free expression.
4. **the right to an effective remedy:** Transparency issues around AI and data-driven systems and liability challenges arising from complex digital value chains create barriers for people seeking redress from potential rights infringements.
5. **the right to social security:** AI, algorithms and digital technologies in welfare systems such as 'digital by default' benefit systems and automated decision-making systems may prevent people from accessing social welfare without discrimination.
6. **the right to work:** workplace uses of AI and digital technologies raise concerns for people's right to fair and equal wages, and rights to non-discrimination, privacy and effective remedies in the workplace. Additionally, platform workers are currently excluded from their right to form and join trade unions under UK employment law.
7. **the right to asylum and freedom of movement:** the increasing deployment of AI and digital technologies in migration contexts such as predictive risk assessments, biometrics, and digital identification systems raise concerns for people's rights to move freely and seek safety from persecution.

Each chapter provides a high-level overview of existing research and evidence, illustrating how digital technologies are reshaping the interpretation and realisation of these rights in practice, and highlighting areas of particular concern in the UK context.

Part II then sets out five recommendations for embedding a coherent, rights-based approach to digital governance in the UK in a way that engages citizens concerns for fundamental human rights in order to rebuild trust and facilitate positive tech policy progress:

1. **A UK Declaration on Digital Rights:** Our primary proposal: Adopt a principle-based declaration, grounded in international human rights law, to provide a clear statement of commitment and a normative framework to guide future digital policy and legislation. An example draft declaration is provided in the Appendix.

Such a declaration must ultimately be translated into practice. The following recommendations should accompany a declarative commitment to digital rights:

2. **Binding and enforceable human rights-based tech regulation:** Introduce robust, cross-sectoral legal frameworks applicable to both public and private actors, with mandatory human rights due diligence and impact assessments throughout the technology lifecycle.
3. **Redlines on unacceptable use cases:** Establish clear prohibitions on the development, deployment, import, and export of technologies that pose unacceptable risks to fundamental rights, without exemptions for public authorities.
4. **Transparency, accountability, and redress:** Strengthen meaningful transparency obligations for technology use, particularly in the public sector, to enable oversight, accountability, and effective access to remedy for individuals and communities.
5. **Meaningful public participation in technology and AI governance:** Ensure that communities affected by digital technologies are meaningfully involved at key points of AI and digital policymaking and governance through deliberative and participatory mechanisms.

Together, these recommendations provide a roadmap for ensuring that technological development in the UK is aligned with human rights, democratic accountability, and the rule of law, and outline an opportunity for working towards a renewed, trusting relationship between citizens and state.

INTRODUCTION

Over the past two decades, digital technology has come to permeate almost every sphere of daily life – from accessing essential public services,[2] to communicating with friends and family members,[3] to navigating workplaces or educational settings,[4] to electoral and democratic processes.[5]

This ongoing period of technological change into the era of AI has brought benefits for some people including enhanced inclusion, greater opportunities for self-expression, and positive outcomes of economic growth.[6] But there are also a growing number of serious concerns about the economic, social, and environmental impacts brought by the emergence of AI and other data-driven digital technologies. These impacts hold potentially serious and wide-ranging implications for our fundamental human rights that feature heavily in the public psyche.[7,8]

Currently, the government has not adequately engaged in the protections needed for society to trust and support this period of transformation. Most recently, public distrust and resistance has featured prominently in debates surrounding the rollout of digital identification in the UK, which is a large digital infrastructure project the government has proposed to simplify people's access to public services.[9] However, it has raised serious concerns about unprecedented state access to personal data enabling large-scale infringement on privacy rights, as well as concerns about fraud, identity exclusion, digital exclusion, and discrimination. The digital ID story demonstrates the political risks around technological changes if the government doesn't engage adequately with citizens on such policies.

At the same time UK government authorities are also increasingly adopting automated surveillance technologies[10] and facial recognition[11] with concerning implications for fundamental rights to privacy, non-discrimination, freedom of movement and asylum, and freedom of expression, information, and assembly. Biased and discriminatory algorithms are also being used to make key decisions about people's lives ranging from welfare distribution,[12] hiring and

2 Big Brother Watch (2021). Poverty Panopticon: the hidden algorithms shaping Britain's welfare state. In C. Van Veen & S. Howes, Big Brother Watch. https://bigbrotherwatch.org.uk/wp-content/uploads/2021/07/Poverty-Panopticon.pdf
3 Brunner, L. (2018). Digital communications and the evolving right to privacy. In Cambridge University Press eBooks (pp. 217–242). https://doi.org/10.1017/9781316838952.010
4 Institute for the Future of Work (2025)). Pissarides Reviews. See: https://www.ifow.org/landing-page/the-pissarides-review
5 Seger, E. and Hancock, J. (2025). Free and Fair: Election law in the age of AI. Demos. https://demos.co.uk/research/free-and-fair-election-law-in-the-age-of-ai/
6 Knight, S. (2025). Tech that Liberates: A new vision for embedding AI in public service reform. Demos.https://demos.co.uk/research/tech-that-liberates-a-new-vision-for-embedding-ai-in-public-service-reform/
7 Modhvadia, R., Sippy, T.,. Field Reid, O., & Margetts, H. (2025) 'How Do People Feel About AI?' (Ada Lovelace Institute and The Alan Turing Institute) https://attitudestoai.uk/
8 Helberger et al. (2025). Governments Want to Ease AI Regulation for Innovation, But Do Citizens Agree? Tech Policy Press. https://www.techpolicy.press/governments-want-to-ease-ai-regulation-for-innovation-but-do-citizens-agree/
9 Seger, E. et al. (2025). Defining Digital ID: Ideating a people centred approach to digitial identification in the UK. Demos. https://demos.co.uk/wp-content/uploads/2026/01/Defining-digital-ID_paper_2026.ac_.pdf
10 Amnesty International UK. (2025). AUTOMATED RACISM: How police data and algorithms code discrimination into policing. https://www.amnesty.org.uk/files/2025-02/Automated%20Racism%20Report%20-%20Amnesty%20International%20UK%20-%202025.pdf
11 Badshah, N. (2025, August 8). Met police to more than double use of live facial recognition. The Guardian.https://www.theguardian.com/technology/2025/jul/31/met-police-to-more-than-double-use-of-live-facial-recognition
12 Big Brother Watch. (2025). Suspicion by Design: What we know about the DWP's algorithmic black box and what it tries to hide. https://bigbrotherwatch.org.uk/wp-content/uploads/2025/07/Suspicion-By-Design-2.pdf

firing decisions,[13] and immigration processes, to criminal legal proceedings such as policing operations, sentencing decisions, or releases.[14] These applications are attracting scrutiny for potential infringements on rights to non-discrimination, freedom of expression, social security, work, freedom of movement and asylum, and fair trial rights.

Meanwhile, the actions of underregulated large technology companies also carry implications for human rights. For example, large online platforms are increasingly influencing what information and narratives people are exposed to online, raising concerns for freedom of expression and the right to access information. Evidence that online narratives are disproportionately amplifying hatred and violence against marginalised communities pose additional concerns for rights to non-discrimination, with tangible offline consequences.[15] Meanwhile this lack of regulation has also led to unprecedented concentration of economic[16] and political power among private companies,[17] creating increasing barriers for states to act in the public interest.

In light of increasingly rapid and socially consequential technological change, this report makes the case for the UK government to make a firm commitment to people's human rights in the digital age. At Demos, we believe that the government must urgently create a new deal between citizens and state to rebuild the breakdown of trust that is fuelling the current democratic emergency; it is a deal in which the state listens to citizens, responds and delivers effectively as a result, and in which citizens trust the state to do so. As part of this deal, a demonstrated commitment to preserving human rights in the face of rapid technological change will be critical.

As the public continues to experience both the opportunities and challenges of technology, it is crucial that people trust that their interests are at the heart of the government's regulatory approach, and that people feel the state is on their side. Currently, we do not have the protections needed for the public to support or sufficiently trust the state's approach to technological adoption to allow even beneficial progress to proceed smoothly. In the absence of adequate legislative frameworks and safeguards ensuring human rights are protected, the government risks facing ongoing resistance to technological transformation which is simultaneously further eroding the trust needed between citizens and state for the well-functioning of democratic society more broadly.

In this report, we gather existing evidence on the impacts of state and corporate uses of technology on people's fundamental human rights, and we urge government to commit to preserving these rights in the face of the challenges we present. One of our core recommendations toward that end is for government to develop and adopt a Declaration on Digital Rights as a human rights-based framework to guide policymaking toward a positive future with technology centred on the wellbeing, autonomy, and dignity of people. We provide an example draft Declaration for further development in the Appendix.

13 Greggwirth. (2025). New study finds AI-enabled anti-Black bias in recruiting - Thomson Reuters Institute. Thomson Reuters Institute. https://www.thomsonreuters.com/en-us/posts/legal/ai-enabled-anti-black-bias/; Worker Info Exchange (2023). Just Eat Report. https://www.workerinfoexchange.org/just-eat-report

14 Statewatch (2025). New Technology, Old Injustice: Data-driven discrimination and profiling in police and prisons in Europe. https://www.statewatch.org/publications/reports-and-books/new-technology-old-injustice-data-driven-discrimination-and-profiling-in-police-and-prisons-in-europe/

15 Perry, H. and Malik, N. (2025). Researching the riots: An evaluation of the efficacy of Community Notes during the 2024 Southport riots. Demos. https://demos.co.uk/research/researching-the-riots-an-evaluation-of-the-efficacy-of-community-notes-during-the-2024-southport-riots/ ; Amnesty International. (2025). Technical explainer on X's recommender system and the 2024 racist riots. https://www.amnesty.org/en/documents/eur45/0618/2025/en/

16 Companies Market Cap (2025). Companies ranked by Market Cap. https://companiesmarketcap.com/gbp/#google_vignette

17 Hao, K. (2025). Empire of AI: Inside the reckless race for total domination. Penguin Books Ltd. ; Srnicek, N. (2026). Silicon empires: The fight for the future of AI. Polity.

GLOBAL PRECEDENT ON DIGITAL RIGHTS

There has been widespread global recognition that rights frameworks and policy intervention are needed to ensure citizens can confidently and equitably enjoy the benefits of our ongoing digital transition, knowing that their fundamental rights will not be compromised for the sake of technological progress. The call is ubiquitous throughout civil society, and several governments and international bodies have initiated responses to uphold fundamental rights in the digital age.

In 2023, the European Commission launched the AI Act with the central aim of "developing a strong regulatory framework based on human rights"[18] and the Biden administration released an Executive Order on Artificial Intelligence to "protect Americans' privacy, advance equity and civil rights".[19] Similar initiatives have been launched in Brazil, the Republic of Korea, Japan, and South Africa.[20] In 2024, the UN has adopted the Global Digital Compact[21] and the Council of Europe adopted the Framework Convention on Artificial Intelligence, becoming the first-ever international legally binding treaty on AI and human rights.[22] The UK is a signatory on both.

DEREGULATORY TREND

However, these efforts are now being met with growing appetite for "digital deregulation".[23] In the global race to maintain technological, geopolitical and economic relevance in the age of AI, regulation is being seen by many as a hindrance to market growth.[24] Technology companies in particular are spearheading efforts to water down protective legislation in the name of progress and economic growth, and governments are prioritising the interests of security actors in regulatory initiatives – such as through the EU AI Act's loopholes that create regulatory exemptions for national security, law enforcement and migration control authorities.[25]

This has played out most prominently through Trump's AI Action Plan, overturning Biden's Executive Order and attempting to place a moratorium on state AI laws, as well as through the Trump administration's tariff threats against countries drafting regulation of US tech companies.[26] In the UK, the GDPR framework has faced amendments weakening its protections through the Data Use and Access Act (2025).[27] Meanwhile the European Commission has proposed 'simplifying' the AI Act by cutting a number of key protections to citizens' rights, allegedly to make it 'workable in practice'.[28]

18 Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)

19 Executive Order No. 14,110, 88 Fed. Reg. 75191 (Oct. 30, 2023). https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence

20 Martins, L. (2025, February 4). Brazil's AI law faces uncertain future as big tech warms to Trump. Tech Policy Press. https://www.techpolicy.press/brazils-ai-law-faces-uncertain-future-as-big-tech-warms-to-trump/ ; Artificial Intelligence Act. (2025, January 9). South Korean AI Basic Law | Artificial Intelligence Act. https://artificialintelligenceact.com/south-korean-ai-basic-law/ ; Japan. (2021). 人工知能関連技術の研究開発及び活用の推進に関する法律. In 法律. https://www.cao.go.jp/houan/pdf/217/217anbun_2.pdf ; Department of Communications and Digital Technologies. (2024). Draft national artificial intelligence policy framework for South Africa [Draft]. https://www.dcdt.gov.za/sa-national-ai-policy-framework/file/338-sa-national-ai-policy-framework.html

21 United Nations Executive Office of the Secretary-General. (2023, May 24). A Global Digital Compact — an Open, Free and Secure Digital Future for All: Our Common Agenda Policy Brief 5. United Nations. https://www.un-ilibrary.org/content/papers/10.18356/27082245-28

22 Council of Europe. (2024). Framework convention on artificial intelligence and human rights, democracy and the rule of law (CETS No. 225).

23 Corporate Europe Observatory (2025). Deregulation Watch. https://corporateeurope.org/en/deregulation-watch

24 Csernatoni, R. (2025, May 20). The EU's AI power play: Between deregulation and innovation. Carnegie Endowment for International Peace. https://carnegieendowment.org/research/2025/05/the-eus-ai-power-play-between-deregulation-and-innovation?lang=en

25 Rodelli, C., & Chander, S. (2025, August 7). One Year On, EU AI Act Collides with New Political Reality. Tech Policy Press. https://www.techpolicy.press/one-year-on-eu-ai-act-collides-with-new-political-reality/

26 Lima-Strong, C. (2025, July 24). Unpacking Trump's AI action plan: gutting rules and speeding Roll-Out. Tech Policy Press. https://www.techpolicy.press/unpacking-trumps-ai-action-plan-gutting-rules-and-speeding-rollout/

27 Statewatch (2025). UK undermining data protection rights and putting EU agreements at risk. https://www.statewatch.org/news/2025/june/uk-undermining-data-protection-rights-and-putting-eu-agreements-at-risk/

28 European Commission. (2025, April 9). AI Continent Action Plan: Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions (COM(2025) 165 final). Publications Office of the European Union. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52025DC0165

The European Commission's 'digital omnibus' package proposes far-reaching amendments that will significantly weaken foundational European tech regulation, including the General Data Protection Regulation (GDPR), protections under the hard-fought-for AI Act, and the e-Privacy framework.[29] This package, and the broader deregulatory agenda, would rollback protections against threats to digital human rights, raising serious concerns for the future of digital rights more broadly.[30] The ripple effects of this EU package could reshape how tech regulation is approached beyond Europe too, including in the UK.[31]

Indeed, currently in the UK, there is no comprehensive cross-sectoral legal framework to govern AI systems. In 2025, the government was set to deliver on its proposal for an AI Bill. However, following repeated delays in presenting the Bill for consultation, projections at the time of writing in January 2026 suggest that the Bill may be abandoned altogether, with the government potentially preferring to rely on existing regulation instead – leaving many technology use cases and sectors unregulated against potential impacts on human rights.[32] Existing regulation that pertains to some AI applications such as the Online Safety Act may help fill the holes but a patchwork of AI relevant legislation across different sectors and regulators risks an incoherent approach to preserving human rights.

Against this broader global trend of tech deregulation, we argue that the UK government must seek to implement robust policy and legislation rooted in coherent human rights-based approaches. To guide the government's regulatory approach to AI and digital technology, we recommend that they commit to a Declaration on Digital Rights, for which we provide a draft in the Appendix.

IN THIS REPORT

As the UK government develops its regulatory approach to AI and digital technologies, this report provides a summary of existing evidence on the impacts of new technologies on our human rights. In doing so, the report encourages the government to place fundamental rights at the centre of its considerations for AI and tech policy to rebuild a trusting relationship between citizen and state and to facilitate positive technological progress.

Part I begins by gathering research and evidence from across the UK digital rights community to evaluate the impacts of AI, algorithms, and data-driven digital technologies on a selection of seven key human rights articles from international frameworks. As policymakers consider future directions for the UK's tech policy regulation, these chapters are intended as high-level summaries for easy consumption.

Part II then turns to recommendations for protecting human rights through existing and forthcoming UK tech policy and legislation. We offer five recommendations to create a new deal for the UK's AI and technology regulation, and rebuild the relationship between citizen and state in the digital age – the first of which is for the government to make a commitment to human rights by developing and adopting a Declaration on Digital Rights to guide future policymaking and legislation. We provide an example draft of a declaration of digital rights for further development in the Appendix.

29 NOYB (2026). EU Commission internal draft would wreck core principles of the GDPR. https://noyb.eu/en/eu-commission-about-wreck-core-principles-gdpr

30 Leufer, D. (2025, November 19). Digital rights are on the chopping block in the European Commission's omnibus. Tech Policy Press. https://www.techpolicy.press/digital-rights-are-on-the-chopping-block-in-the-european-commissions-omnibus/ ; EDRi et al. (2025, November 17). The EU must uphold hard-won protections for digital human rights. https://edri.org/wp-content/uploads/2025/11/The-EU-must-uphold-hard-won-protections-for-digital-human-rights.pdf

31 Jahangir, R. (2025, November 10). EU set the global standard on privacy and AI. Now it's pulling back. Tech Policy Press. https://www.techpolicy.press/eu-set-the-global-standard-on-privacy-and-ai-now-its-pulling-back/

32 Bristow, T. (2025, December 23). How the UK fell out of love with an AI bill. POLITICO. https://www.politico.eu/article/how-labour-fell-out-love-with-ai-bill-peter-kyle/

This report was developed between May 2025 and January 2026 based on a literature review on the digital rights movement, legal and policy analysis by the Oxford Martin School of AI Governance on global AI policy trends, and interviews with over 40 experts in technology and human rights in the UK and internationally.

PART 1: HUMAN RIGHTS IN THE DIGITAL ERA

ASSESSING THE IMPACTS OF AI & DATA-DRIVEN TECHNOLOGIES

The following chapters gather evidence from across digital rights literature and expert interviews illustrating the real and potential impacts of new and emerging digital technologies in the UK on a set of seven human rights:

1. **Right to equality & non-discrimination**
2. **Right to privacy**
3. **Right to freedom of expression, information & assembly**
4. **Right to an effective remedy**
5. **Right to social security**
6. **Right to work**
7. **Right to asylum & freedom of movement**

This is by no means an exhaustive list of human rights, but each was selected for its particular relevance in a rapidly changing digital world. These rights are rooted in the Universal Declaration of Human Rights (UDHR), the International Covenant on Civil and Political Rights (ICCPR), and the International Covenant on Economic, Social and Cultural Rights (ICESCR). Our analysis also draws on Convention 108 in recognition of the right to data protection as a fundamental human right, the United Nations Convention on the Rights of the Child (UNCRC) in recognition of the specific impacts of AI and digital technology on children and young people, and the UN Business and Human Rights Guiding Principles in recognition of the outsized role of technology companies in digital rights abuses.

TABLE 1
ACRONYMS

UDHR	Universal Declaration of Human Rights (1948)
ICCPR	International Covenant on Civic and Political Rights (1966)
ICESCR	International Covenant on Economic, Social and Cultural Rights (1966)
UNCRC	United Nations Convention on the Rights of the Child (1989)
ECHR	European Convention on Human Rights (1950)
HRA	Human Rights Act (1998)
UK GDPR	UK General Data Protection Regulation (2018)

Each chapter analyses how the right has been interpreted in the digital age and examines the implications of AI and digital technologies for its realisation. Although these rights emerged in an offline context, their underlying principles remain directly applicable to the digital environment, providing globally recognised standards for safeguarding democracy, dignity, equality, and accountability. In the context of declining public trust in government, and democratic erosion, anchoring the regulation of digital technologies in human rights frameworks is essential to reaffirm states' obligations to the public, democracy, and the rule of law, while ensuring continuity through established legal precedent and international consensus as societies transition from offline to online.

CHAPTER 1
RIGHT TO EQUALITY AND NON-DISCRIMINATION

The right to equality and non-discrimination as articulated in the UDHR states:

Article 1: *All human beings are born free and equal in dignity and rights.*

Article 7: *All are equal before the law and are entitled without any discrimination to equal protection of the law. All are entitled to equal protection against any discrimination in violation of this Declaration and against any incitement to such discrimination.*[33]

A similar articulation is offered by the ICCPR:

Article 26: *All persons are equal before the law and are entitled without any discrimination to the equal protection of the law. In this respect, the law shall prohibit any discrimination and guarantee to all persons equal and effective protection against discrimination on any ground such as race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status.*[34]

1.1 ABOUT THE RIGHT TO EQUALITY AND NON-DISCRIMINATION

The principles of equality and non-discrimination are foundational to the rule of law. The right emerged from early philosophical writings which emphasised human dignity and this was then enshrined in the very first article of the UDHR which highlights the equality of all people.

Equality and non-discrimination is an 'enabling right' which means that, without it, other rights cannot be realised. When inequality exists, minority groups can be excluded, human dignity is impaired and there is a barrier to engaging in economic, social and political life. This can cause and perpetuate poverty, restrict life changes, exacerbate health problems and foster violence and tension.

Progress has been made in ensuring equality and non-discrimination through international legislation such as the Convention of the Elimination of Discrimination against Women and Convention on the Rights of Persons with Disabilities. At the domestic level the UK employs the

33 United Nations General Assembly. (1948). Universal declaration of human rights (217 [III] A). https://www.un.org/en/about-us/universal-declaration-of-human-rights

34 United Nations General Assembly. (1966, December 16). International Covenant on Civil and Political Rights. Treaty Series, 999, 171. https://www.refworld.org/legal/agreements/unga/1966/en/17703

Equality Act 2010. However continued progress is necessary, especially as the proliferation of digital technologies presents new implications for the preservation of human rights.

1.2 IMPACTS OF DIGITAL TECHNOLOGY ON THE RIGHT TO NON-DISCRIMINATION

There are clear implications for the right to non-discrimination posed by new and emerging digital technologies. This section will attend specifically to how automated decision-making systems and social media platforms have a tendency to further discrimination against already marginalised groups and how this interacts with the right to non-discrimination.

1.2.1 Algorithmic discrimination:

Algorithmic discrimination refers to situations where an algorithm systematically produces unfair or unequal outcomes for certain groups of people, often along lines such as race, gender, age, or immigration status.[35] It is the result of the algorithms being trained on historical data that encodes existing social inequalities and biases. For example, if human decision-making has led to immigrants being disproportionately denied home mortgages, that pattern will appear in the historical lending data. When algorithms are trained on such biased data without careful correction, they tend to learn and reproduce those same patterns, embedding past inequities into their outputs (e.g. predictions or automated decisions). When biased outputs are used to inform decisions, the biases are perpetuated and reinforced in the real world.

Algorithmic discrimination has been a particular concern where algorithms have been used to replace or support decision-making in the criminal justice system, for example, for 'predicting' an individual's likelihood of committing a crime and for'predicting' which areas in a city a crime will be most likely to take place.[36] Built using swathes of historic crime data, in which racially minoritised groups are known to be over-represented, crime prediction algorithms, commonly known as 'predictive policing' have been widely criticised by policymakers, politicians, and human rights experts for perpetuating and reinforcing existing racial discrimination within the criminal justice system, which can lead to the targeted policing of racially minoritised groups through technological decision-making or decision-support.[37]

In some cases, the use of crime prediction algorithms has been ruled unlawful because of racial discrimination, such as with London Metropolitan Police's Gangs Matrix.[38] The database contained personal information of people perceived to be in a gang or likely to commit violence, and assigned individuals with an automated 'harm score'. The 'harm score' was generated using a crime prediction algorithm drawing on individuals' data from the database to make predictions about their alleged risk of harm.[39] Research found that racially minoritised groups were significantly overrepresented on the database: 78% of people on the Matrix were Black males and 15% were children, with some as young as 12.[40] In 2018, the Information Commissioner ruled that the Matrix consistently breached data protection laws because of racial disproportionality, forcing the Met to concede that their operation of the Matrix was unlawful.[41]

35 Sombetzki, P. (2026). What is algorithmic discrimination? AlgorithmWatch. https://algorithmwatch.org/en/what-is-algorithmic-discrimination/
36 Ferris, G., Min, B., Nayak-Oliver, M.. (2021). AUTOMATING INJUSTICE: THE USE OF ARTIFICIAL INTELLIGENCE & AUTOMATED DECISION-MAKING SYSTEMS IN CRIMINAL JUSTICE IN EUROPE. Fair Trials. https://www.fairtrials.org/app/uploads/2021/11/Automating_Injustice.pdf
37 Amnesty International UK. (2025). AUTOMATED RACISM: How police data and algorithms code discrimination into policing. https://www.amnesty.org.uk/files/2025-02/Automated%20Racism%20Report%20-%20Amnesty%20International%20UK%20-%202025.pdf
38 Liberty. (2022, November 11). Met to overhaul 'racist' Gangs Matrix after landmark legal challenge - Liberty. https://www.libertyhumanrights.org.uk/issue/met-to-overhaul-racist-gangs-matrix-after-landmark-legal-challenge/
39 Amnesty International. (2018). TRAPPED IN THE MATRIX. In Amnesty International United Kingdom Section. https://www.amnesty.org.uk/files/reports/Trapped%20in%20the%20Matrix%20Amnesty%20report.pdf
40 Ibid.
41 Mohdin, A. (2018, November 16). Met's "gang matrix" breached data laws, investigation finds. The Guardian. https://www.theguardian.com/uk-news/2018/nov/16/met-police-gang-matrix-breached-data-laws-investigation-finds#:~:text=The%20ICO%20investigation%20found%20that%20the%20gang%20matrix%20failed%20to,shared%20with%20other%20public%20bodies.

While the Gangs Matrix has since been overhauled, there are many similar crime prediction systems in operation by criminal justice authorities across the UK. Indeed, a recent report by Amnesty International UK titled 'Automated Racism' found that almost three quarters of all UK police forces have used or are using either geographic or individual crime prediction systems.[42] Notable examples of crime prediction algorithms by UK criminal justice authorities include: the Ministry of Justice's OASys recidivism prediction algorithm used for every person entering the criminal justice system, and which Statewatch found profiles over 1,300 people daily using sensitive personal information such as criminal history, emotional wellbeing, income, and employment;[43] and a geographic hotspot prediction algorithm under the Home Office's Grip programme that is used by 20 police forces across the UK, and which Amnesty UK found "reinforced and contributed to racial profiling and racist policing" following demographic analysis.[44]

Because of the mounting evidence that crime prediction algorithms reinforce existing racial discrimination within the criminal justice system, there are many calls among policymakers and politicians for a ban on crime prediction systems. In 2025, a cross-party group of eight MPs, led by Green Party MP Sian Berry, tabled an amendment to the Crime and Policing Bill that would specifically prohibit the use of automated decision-making (ADM), profiling and artificial intelligence (AI) for the purpose of making risk assessments about the likelihood of groups or people committing criminal offences.[45] Indeed, the EU AI includes a similar clause that places a prohibition on all individual crime prediction systems.[46]

Similar issues of algorithmic discrimination are also well-documented in relation to facial recognition which have been proven to entrench racial stereotypes and perpetuate inequalities. Algorithms that are deployed to identify individuals are predominantly trained on white, male, European faces. Consequently they have a statistically less chance of accurately identifying faces of other ethnic backgrounds.[47] In some cases, facial recognition algorithms misclassified black women nearly 35% of the time while almost always correctly identifying white men.[48] This is a particular issue when police forces use facial recognition technology to identify potential criminals. A case in the United States saw police forces in New York city used facial recognition technology to arrest a suspect for sexual assault. Due to errors in the technical system, they falsely arrested and jailed a black man for two nights despite mobile phone location data showing that the suspect was miles away from the crime scene at the time of the crime.[49] Meanwhile London's Metropolitan Police have doubled their deployment of facial recognition vans up to 10 uses per week while Cardiff police use live facial recognition during concerts and sporting events.

1.2.2 Discrimination on online platforms

The emergence of social media platforms, which effectively function as the online public sphere, has opened up new spaces where existing discrimination and oppression in society

42 Amnesty International UK. (2025). AUTOMATED RACISM: How police data and algorithms code discrimination into policing. https://www.amnesty.org.uk/files/2025-02/Automated%20Racism%20Report%20-%20Amnesty%20International%20UK%20-%202025.pdf
43 Statewatch (2025, April 9). Over 1,300 people profiled daily by Ministry of Justice AI system to 'predict' re-offending risk. https://www.statewatch.org/news/2025/april/uk-over-1-300-people-profiled-daily-by-ministry-of-justice-ai-system-to-predict-re-offending-risk/
44 Amnesty International UK. (2025). AUTOMATED RACISM: How police data and algorithms code discrimination into policing. https://www.amnesty.org.uk/files/2025-02/Automated%20Racism%20Report%20-%20Amnesty%20International%20UK%20-%202025.pdf
45 Skelton, S. K. (2025, June 27). MPs propose ban on predictive policing. ComputerWeekly.com. https://www.computerweekly.com/news/366626658/MPs-propose-ban-on-predictive-policing
46 Fair Trials (2023, December 11). Partial ban on 'predictive' policing and crime prediction systems included in final EU AI Act. https://www.fairtrials.org/articles/news/partial-ban-on-predictive-policing-included-in-final-eu-ai-act/
47 Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. Proceedings of Machine Learning Research, 81, 77-91. https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf
48 Crockford, K. (2023, July 17). How is Face Recognition Surveillance Technology Racist? | ACLU. American Civil Liberties Union. https://www.aclu.org/news/privacy-technology/how-is-face-recognition-surveillance-technology-racist
49 Cranmore, C. (2025, August 27). Man's wrongful arrest puts NYPD's use of facial recognition tech under scrutiny. ABC7 New York. https://abc7ny.com/post/man-falsely-jailed-nypds-facial-recognition-surveillance-tech-failed/17664671/

can proliferate through online means. This section considers some of the ways that online platform design plays a role in reinforcing existing discrimination, including racism, misogyny, ableism, and homophobia through content moderation systems and recommender algorithms. The simultaneous failure to moderate discriminatory content and the amplification of this discriminatory content through recommender algorithms has led to offline harm and violence against communities already experiencing marginalisation.

Firstly, research suggests that content moderation systems on social media platforms are often ineffective at moderating discriminatory online content. A prime example of this is during the 2024 Southport riots in the UK, when social media platforms, especially X, failed to effectively moderate content inciting racial hatred and violence. Our research at Demos looked at posts on X over the course of the riots finding more than one in five posts were directly threatening to racialised groups. As Naema Malik, Researcher at Demos, writes for the Big Issue: "One post that falsely claimed the attacker was Muslim received 1.5 million views. Another, alleging he was an "illegal immigrant", reached nearly 1.3 million. More than half of these harmful posts targeted migrants, 36% focused on Muslims and a third were explicitly xenophobic or racist."[50] Our research subsequently evaluated the effectiveness of X's Community Notes moderation system, finding that it was fundamentally unfit for purpose at moderating discriminatory and racist content during the riots. Community Notes is a feature allowing users to collaboratively add context and corrections to potentially misleading posts.[51] Our research found that during the riots less than 5% of community notes proposed were published and, of the notes that did go live, the time it took to publish them failed to mitigate against people seeing, and acting on, false and harmful content.[52] The ineffectiveness of X's moderation system during the Southport riots is an example of the ways that failures of content moderation systems on social media can allow discriminatory content to continue proliferating on online platforms and, as during the riots, lead to offline violence against marginalised communities.

Secondly, evidence from digital rights advocates suggests that the recommender algorithms on social media platforms disproportionately amplify discrimination. The business model of social media platforms relies on the sale of targeted advertisements and, in order to produce this revenue, platforms must ensure that users stay on the platform for as long as possible, leading platforms to implement systems and processes that first and foremost prioritise engagement – with research showing that discriminatory or hateful content generates the most engagement.[53] For example, Amnesty International's research has demonstrated that: X's recommender algorithm prioritised racist content during the Southport riots and amplified hate speech towards the LGBTI community in Poland,[54] while Facebook's content algorithm amplified racist claims about the Rohingya minority in Myanmar.[55] In all of these examples, Amnesty found that the recommender algorithms on social media platforms contributed to offline violence against marginalised communities, demonstrating the real-life consequences of online discrimination.

50 Malik, N. (2025, August 28). From Southport to Epping, social media's failure to act is fuelling racist violence. Big Issue. https://www.bigissue.com/opinion/southport-epping-social-media-racism-violence/

51 Perry, H. and Malik, N. (2025). Researching the riots: An evaluation of the efficacy of Community Notes during the 2024 Southport riots. Demos. https://demos.co.uk/research/researching-the-riots-an-evaluation-of-the-efficacy-of-community-notes-during-the-2024-southport-riots/

52 Ibid.

53 Amnesty International. (2025). Technical explainer on X's recommender system and the 2024 racist riots. https://www.amnesty.org/en/documents/eur45/0618/2025/en/

54 Ibid.

55 Amnesty International. (2023). Myanmar: Facebook's systems promoted violence against Rohingya; Meta owes reparations – new report. https://www.amnesty.org/en/latest/news/2022/09/myanmar-facebooks-systems-promoted-violence-against-rohingya-meta-owes-reparations-new-report/

CHAPTER 1 SUMMARY POINTS:

- The right to non-discrimination highlights the equality of all people. It is an 'enabling right', meaning it is essential for the fulfillment of all other human rights.
- AI and digital technology use cases that may reinforce discrimination against marginalised groups include: algorithms, automated decision-making systems, and content moderation and recommender algorithms on social media platforms.
- Automated decision-making systems can encode discrimination when training data under- or overrepresents certain groups of people. This can lead to negative real-life outcomes for people.
- Content moderation systems on social media platforms can fail to effectively moderate discriminatory content.
- Research suggests that recommender algorithms on social media platforms amplify inflammatory content which often equates to discriminatory content.
- Protecting the right to non-discrimination in the digital age therefore requires robust procedural safeguards at all levels of technology governance and deployment. In some instances, prohibitions on certain use cases are necessary.

CHAPTER 2
RIGHT TO PRIVACY IN THE DIGITAL AGE

The right to privacy as articulated in Article 12 of the UDHR and Article 17 of the ICCPR states: "*No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks.*"[56]

As well as article 12 of the UDHR and article 17 of the ICCPR, the right to privacy is also enshrined in article 8 of the ECHR, and brought into UK law by the Human Rights Act. Over 185 national constitutions mention the right to privacy.

2.1 ABOUT THE RIGHT TO PRIVACY

The right to privacy is a foundational right for protecting our lives against interference and intrusion from both state and private actors, both online and offline. It allows us to have the space we need to exercise our autonomy, express ourselves freely and without judgement, to form our own thoughts and opinions, and to develop our sense of identity free from external control. It protects our family life, our home, and our communications against unchecked state and corporate power or control.

The right to privacy is not, however, an absolute right, meaning that it may be subject to certain restrictions, but only if they meet a stringent three-part test. The restrictions must be:

1. provided by law (which must be formulated with enough precision to enable an individual to regulate their conduct accordingly);
2. demonstrably necessary and proportionate (using the least restrictive measure to achieve the specified purpose);
3. for the purpose of protecting specified public interests (such as national security) or the rights or reputations of others.

56 United Nations General Assembly. (1948). Universal declaration of human rights (217 [III] A). https://www.un.org/en/about-us/universal-declaration-of-human-rights ; United Nations General Assembly. (1966, December 16). International Covenant on Civil and Political Rights. Treaty Series, 999, 171. https://www.refworld.org/legal/agreements/unga/1966/en/17703

Today, the right to privacy extends to the vast landscape of digital data that underpins the digital transformation – including the processing of enormous quantities of information, often including sensitive personal information, for AI and algorithms, cloud computing, or data analytics.

Following the digitisation of society, the right to privacy was expanded to encompass data protection. The Council of Europe's 1981 treaty "The Convention for the Protection of Individuals with Regard to the Automatic Processing of Personal Data" (known as Convention 108) was the root treaty that spawned the first EU-wide data protection laws. Globally, there are 132 jurisdictions that have data privacy laws covering similar concerns in different ways, sometimes using different terms. The right to data protection, as articulated in Convention 108, includes the following seven principles:[57]

- **Lawfulness, fairness and transparency:** Everyone has a right to lawful and fair processing of personal data, with clear and easily understandable information about that processing.
- **Purpose limitation:** Personal data should only be collected for specified, explicit, and legitimate purposes.
- **Data minimisation:** Processing of personal data must be adequate, relevant, and limited to what is necessary in relation to the purposes for which they are processed.
- **Accuracy:** Data controllers must ensure that personal data is accurate and take every reasonable step to erase or rectify inaccurate data without delay.
- **Storage limitation:** Personal data permitting identification of data subjects should not be kept longer than necessary, with established time limits.
- **Integrity and confidentiality:** Personal data should be processed in a manner that ensures appropriate security, including protection against unauthorised access.
- **Accountability:** Data controllers must take responsibility and demonstrate their compliance with the GDPR.

The EU's General Data Protection Regulation (GDPR) (2016) built on the foundational principles of Convention 108 to establish comprehensive regulation for the processing of personal data and grant data subjects eight primary rights over their data: right to be informed; right of access; right to rectification; right to erasure; right to restrict processing; right to data portability; right to object; and rights in relation to automated decision-making and profiling.

The rights to privacy and data protection also intertwine with other rights, where infringements on the rights to privacy and data protection can also lead to violations of other other rights such as the right to non-discrimination, and freedom of expression and association.

2.2 IMPACTS OF DIGITAL TECHNOLOGY ON PRIVACY

While the GDPR has been rightly celebrated as improving data protection rights, it is also considered insufficient for fully protecting people's right to privacy in the digital age for a number of factors: data protection authorities often lack sufficient resourcing to adequately enforce the GDPR; technological developments mean that certain use cases are not adequately regulated; exemptions for public authorities such as national security and law enforcement leave people vulnerable to many high-risk use cases; recent backtracking to the GDPR framework; and the onus placed on the individual by the GDPR is often inadequate to address privacy concerns arising from AI and new technologies.

57 Data Protection Commission (no date) 'Principles of Data Protection' Data Protection Commission https://www.dataprotection.ie/en/individuals/data-protection-basics/principles-data-protection

This section covers some present and emerging examples where the right to privacy is threatened by state and corporate uses of AI, digital technology, and data – whether through surveillance technologies such as facial recognition and crime prediction systems, the ongoing surge in generative AI, or online platforms' use of targeted advertising.

The examples listed demonstrate that with access to vast caches of information, states and corporations can acquire unprecedented influence and control over people's lives, despite existing regulatory protections – for example, to target people with hyperpersonalised content (e.g. online advertisements and political propaganda), to employ automated decision making tools to make determinations about important aspects of people's lives (e.g. regarding criminal proceedings, welfare approvals, and immigration status), to surveil the movement and behaviors of individuals and populations at scales, or to simply sell the data and its cornucopia of potential uses on to third parties without user knowledge or consent and without guarantee of responsible use.

More so, the examples demonstrate that privacy violations disproportionately impact some people and communities more than others such as journalists working to expose corrupt government activities, human rights campaigners and activists, and already marginalised groups such as migrants, people of colour, and people with disabilities.

The section is by no means exhaustive. Instead, it gathers a selection of pressing examples that have raised particular concerns among the digital rights community, despite existing regulatory protections.

2.2.1 Impacts on the right to privacy of AI and digital technology used by state actors

The collection and processing of personal data is becoming increasingly embedded in state activities – from a growing network of databases, data-sharing methods, and increasingly sophisticated tech tools for processing data. Of particular note are the present government's commitments to embedding technology in public service reform, such as through the increasing deployment of facial recognition technologies, the use of AI and algorithms for automated decision-making, and the recent drive to establish a digital ID program, all of which will rely upon extensive personal data processing. Consequently there is growing concern among many digital rights groups, civil society organisations, and the public about the use of data and digital technology for surveillance and monitoring purposes, presenting threats to peoples' right to privacy.

Table 1 presents a wide sampling of new and emerging digital technologies used by the state that have raised concerns about privacy infringements because in many cases they are deployed without a sufficient legal basis, involve disproportionate data processing, are unnecessary for their purported aims, and carry high risks of discrimination.

In response to these varying privacy concerns, digital rights activists and organisations have been undertaking sustained campaigns and advocacy initiatives to call for greater regulation around government uses of technology that entails widespread processing of personal and sensitive data. In some cases, the digital rights community has called for outright prohibitions around certain use cases. For example: Big Brother Watch has led a long-standing campaign for a prohibition on LFR in public spaces, saying that the technology depends on disproportionate and unnecessary data processing which raises serious concerns for discrimination, particularly against racialised people; Amnesty UK and Open Rights Group have called for a prohibition on crime prediction systems, because the technology has been proven to reinforce structural racism, and involves the indiscriminate processing of sensitive data about people and their criminal backgrounds; while a recent civil society-wide campaign on digital ID highlighted the indiscriminate and disproportionate processing of people's data as well as security risks.

As an absolute minimum requirement, there are calls for greater transparency, and robust regulatory frameworks and enforcement, giving data subjects meaningful avenues for consent and redress. The table below summarises a selection of these campaigns and advocacy initiatives.

TABLE 2
STATE USES OF AI AND DATA-DRIVEN TOOLS AND ASSOCIATED PRIVACY THREATS

HOW DOES IT WORK?	PRIVACY THREAT	IMPACTS/EXAMPLES	CAMPAIGNS
Live facial recognition (LFR)			
Allows law enforcement authorities to scan the faces of passersby, compare their biometric 'faceprint' against 'watchlists' and generate alerts for possible 'matches' for police to follow up on. LFR is usually deployed at busy locations such as shopping streets, large events, and protests.	Involves the indiscriminate and widescale extraction and retention of unique biometric identifiers about people in public spaces. The technology is currently unregulated. Also impacts people's rights to non-discrimination, and freedom of expression and association.	In July 2025, London's Metropolitan Police doubled its usage of LFR up to 10 deployments per week, including deploying LFR at Notting Hill Carnival. Fixed facial recognition is increasingly used too; in Cardiff there is a ring of 11 fixed cameras that are switched on during busy periods, such as sports events and concerts. In 2020, the English Court of Appeal ruled that South Wales police's use of LFR was unlawful. Despite this ruling, LFR is still deployed extensively. LFR can lead to people being wrongfully stopped, searched, questioned, and arrested, such as in the case of anti-knife crime campaigner Shaun Thompson.	In 2023, 65 cross-party MPs and peers called for an "immediate stop" on LFR in public spaces. The call was backed by 31 UK rights groups.[58] In 2024, the UN HRC said that the UK should "end the use of facial recognition" by law enforcement.[59] Previously, 180 global experts have called for a stop on LFR in public spaces.[60]

58 Big Brother Watch Team (2023). '65 parliamentarians call for 'immediate stop' to live recognition surveillance'. Big Brother Watch https://bigbrotherwatch.org.uk/press-releases/65-parliamentarians-call-for-immediate-stop-to-live-facial-recognition-surveillance/
59 Newson (2024). 'UN standards on the use of surveillance technology at protests'. House of Lords Library https://lordslibrary.parliament.uk/un-standards-on-the-use-of-surveillance-technology-at-protests/
60 Big Brother Watch Team (2023). '180+ tech experts call for global stop to facial recognition surveillance'. Big Brother Watch https://bigbrotherwatch.org.uk/press-releases/180-tech-experts-call-for-global-stop-to-facial-recognition-surveillance/

HOW DOES IT WORK?	PRIVACY THREAT	IMPACTS/EXAMPLES	CAMPAIGNS
Digital ID			
Digital ID systems allow governments and other actors to collect, store, and share records which can be used to verify the identity of individuals. Common data types used include personal information such as names and birth dates, biometric data, and government-issued ID codes.	Enables the persistent surveillance of individuals' behaviours by associating data generated through activities with their unique identity. This privacy risk is exacerbated when digital ID verification is made a requirement for access to goods and services, such as purchases or access to welfare. Depending on the system's design, there may be an additional privacy risk if the system undergoes a cybersecurity breach that exposes individuals' sensitive personal data. Affects rights to privacy and may enable violations of the right to non-discrimination. If a system is mandatory for accessing essential goods and services or voting, those unable or unwilling to use the system may experience violations of other rights, including the right to social security, education, participation in civic life, and political expression.	India's 'Aadhaar' digital ID system was launched in 2016 and is the largest system of its kind in the world, with 1,427,687,248 Indians enrolled as of October 2025. Aadhaar relies on biometric data and is a prerequisite for accessing essential government services.[61] Aadhaar has been the repeated subject of criticisms for privacy violations and negative impacts on digital inclusion. In 2018, a case was brought to India's Supreme Court over Aadhaar's privacy and security impacts, which resulted in significant aspects of the programme being shut down. Following the UK government's announcement of its intention to introduce a 'BritCard' digital ID system human rights organisations such as Liberty have raised concerns about negative impacts on privacy and digital inclusion.	Organisations in the UK that are running campaigns highlighting the privacy risks of digital ID include Big Brother Watch, Liberty, and the Open Rights Group.[62]

61 Aadhaar dashboard. https://uidai.gov.in/aadhaar_dashboard/; Chandran, R. (2017, 13 December). India's digital ID sparks debate over human right to personal data. Reuters. https://www.reuters.com/article/world/indias-digital-id-sparks-debate-over-human-right-to-personal-data-idUSKBN1E71DA/ ; Lanka, S. N. (2025, May 26). When KYC becomes a barrier: Supreme Court's stand for digital inclusion. Internet Freedom Foundation (IFF). https://internetfreedom.in/when-kyc-becomes-a-barrier-supreme-courts-stand-for-digital-inclusion/ ; Liberty. (2025, September 26). Compulsory digital ID will exclude some of the most marginalised members of society. https://www.libertyhumanrights.org.uk/issue/compulsory-digital-id-will-exclude-some-of-the-most-marginalised-members-of-society/ ; Big Brother Watch. (2025). Checkpoint Britain: The dangers of digital ID and why privacy must be protected. https://bigbrotherwatch.org.uk/wp-content/uploads/2025/09/Checkpoint-Britain.pdf

62 Big Brother Watch. (2025b, September 23). No2DigitalID. https://bigbrotherwatch.org.uk/campaigns/no2digitalid/ ; Liberty. (2025b, September 26). LIBERTY'S POSITION ON DIGITAL ID. https://www.libertyhumanrights.org.uk/issue/digital-id-liberty-position/ ; Open Rights Group (2025). ID Cards: UK Risks Sleepwalking into a Pre-Crime State. https://www.openrightsgroup.org/press-releases/id-cards-uk-risks-sleeping-walking-into-pre-crime-state/

HOW DOES IT WORK?	PRIVACY THREAT	IMPACTS/EXAMPLES	CAMPAIGNS
Crime prediction systems			
AI and algorithmic systems used in criminal legal settings compare data points about individuals or places against historic crime databases to make 'predictions' about the risk of crime by a person or place. They are an example of automated recommendation-making systems.	Relies on enormous databases holding sensitive personal information about people (eg. criminal history, ethnicity, age, associations etc), with high risks for racial disproportionality. Affects rights to non-discrimination, freedom of expression and association, fair trial, and right to liberty.	As of 2025, at least 75% of all UK police forces are using crime prediction algorithms, either geographic or individual. A notable example is the, now scrapped, 'Gangs Matrix' used by the Met police which evaluates people's alleged risk of being in a gang. Data used included: ethnicity, criminal justice data, uncorroborated police intelligence, social media posts. A disproportionate amount of the database, 78%, were Black males, showing clear evidence of racial discrimination. Recorded consequences included: increased stop and search, immigration action, and school exclusions; and disruptions to prison licence conditions, benefits entitlements, and housing arrangements, including evictions. The Matrix has been discontinued after a legal challenge, but it has been replaced by the Violence Harm Assessment (VHA). Other crime prediction systems include: the Ministry of Justice's OASys; Essex Police's Knife Crime and Violence Model; Avon and Somerset Police's Qlik Sense; and Greater Manchester Police's XCalibre system.	Over 30 UK civil society organisations have called for an outright prohibition on crime prediction systems, including Amnesty UK, Open Rights Group, Big Brother Watch, Public Law Project, Statewatch and Liberty.[63] Calls for a prohibition have been supported by a cross-party group of MPs including Sian Berry, Zarah Sultana, Ellie Chowns, Richard Burgon, and Clive Lewis.[64] In the EU, 54 civil society organisations called for a prohibition on crime prediction in the EU AI Act.[65] In June 2023, the European Commission adopted these recommendations, and implemented a prohibition on individual crime prediction systems in the Act.[66]

63 Statewatch (2025). 'Law enforcement use of automated decision making' [letter] State Watch https://www.statewatch.org/media/4874/uk-law-enforcement-adm-letter-21-3-25.pdf
64 Skelton (2025). 'MPs propose ban on predictive policing'. Computer Weekly https://www.computerweekly.com/news/366626658/MPs-propose-ban-on-predictive-policing
65 Fair Trials and EDRi (no date) 'Civil Society, rights groups calls on the EU to prohibit predictive and profiling AI systems in law enforcement and criminal justice' Fair Trails. https://www.fairtrials.org/app/uploads/2022/03/Prohibit-predictive-and-profiling-AI-systems-in-law-enforcement-and-criminal-justice-January-20239828ca9d610e35808e4baf8ff26014fb406cb1bfdffa979b037f53a387896e87.pdf
66 Fair Trials (2023). 'Partial ban on predictive policing included in final EU AI Act'. Fair Trials https://www.fairtrials.org/articles/news/partial-ban-on-predictive-policing-included-in-final-eu-ai-act/#:~:text=After%20months%20of%20negotiations%2C%20the,been%20campaigning%20for%20since%202021.

HOW DOES IT WORK?	PRIVACY THREAT	IMPACTS/EXAMPLES	CAMPAIGNS
Social media monitoring			
Involves the monitoring, extraction and retention of public and private social media content by government authorities – both by automated and manual means. Several UK government bodies undertake social media monitoring, including: the DWP, DfE, DCMS, DfH, DEFRA, and DBT.	Involves the widescale monitoring of private correspondences – without an adequate legal basis, or scrutiny on its necessity or proportionality. Affects rights to non-discrimination, freedom of expression and association, and freedom of thought, conscience and religion.	In Catt vs United Kingdom, the National Domestic Extremism Unit was found to hold personal details of 9,000 campaigners, amounting to privacy violations. As of 2020, 60% of local authorities were undertaking social media monitoring. Clearview AI has scraped up to 30 billion facial images, including those of UK residents, to train its facial recognition technology. In the case of Raza v The City of New York, social media monitoring was used to target minority ethnic groups.	Privacy International has previously campaigned for adequate safeguards, effective oversight, and meaningful accountability for social media monitoring.
Bulk communication interceptions			
Interception of communications, both content and metadata, transiting undersea fiber optic cables.	Interferes with our right to private correspondence. Affects rights to freedom of expression and association, and freedom of thought, conscience and religion.	In Big Brother Watch and Others vs UK (2021), 10 NGOs challenged the UK government's bulk interception of internet traffic from the UK to US via undersea cables. The ECtHR ruled that metadata could be as revealing as the content of communications themselves, marking an expansion in understandings of privacy.	-

HOW DOES IT WORK?	PRIVACY THREAT	IMPACTS/EXAMPLES	CAMPAIGNS
Weakening encryption			
Encryption is a method of scrambling data so that only those with the correct key may access and understand it. Weakening encryption allows for unauthorised access to communications, systems, and devices. This access can be used to gather sensitive private data, intercept conversations, or to facilitate further cybersecurity breaches. Weakened encryption can also reduce the overall security of a device or system, allowing for other malicious third parties to gain access after the initial breach.	Enables violations of the right to privacy, family, home life, and correspondence. By enabling other abuses, such as bulk communication interceptions, these violations can affect rights to freedom of expression and association, and freedom of thought and conscience.	In February 2025, the Washington Post revealed that the UK government had requested that Apple create a backdoor in its encrypted Advanced Data Protection (ADP) setting for iOS devices, under the Investigatory Powers Act 2016.[67] Apple responded by withdrawing ADP from the UK. A legal battle ensued where Apple, Privacy International and other complainants argued that the request would violate iOS users' security and right to privacy. While it was said that the complaint would go to tribunal in 2026, subsequent reporting has suggested that the UK government has backed down.	The Open Rights Group's 'Save Encryption' and 'Practice Safe Text' campaigns call for the protection of encryption in the UK. The campaigns target provisions in the Online Safety Act and Investigatory Powers Act which could give the government the power to request access to encrypted services.[68]

67 Menn, J. (2025, February 7). U.K. orders Apple to let it spy on users' encrypted accounts. The Washington Post. https://www.washingtonpost.com/technology/2025/02/07/apple-encryption-backdoor-uk/ ; Kleinman, Z. (2025, March 4). Apple takes legal action in UK data privacy row. BBC News. https://www.bbc.co.uk/news/articles/c8rkpv50x01o ; Kleinman, Z. (2025b, August 19). UK backs down in Apple privacy row, US says. BBC News. https://www.bbc.co.uk/news/articles/cdj2m3rrk74o

68 Open Rights Group (2025). Save Encryption. https://www.openrightsgroup.org/campaign/save-encryption/

2.2.2 Impacts on the right to privacy of data and digital technology used by private actors

The previous section presents a selection of pressing examples where state data processing raises concerns for the right to privacy. This section will cover examples of data processing by private actors that present risks to users' privacy rights. In particular, the section summarises some of the privacy and data protection concerns surrounding targeted advertising and generative AI which rely on the widespread collection of personal data that often undermine key principles for data processing such as consent and user control. The state has a responsibility to respect, protect and fulfill the rights of those in its jurisdiction even when the right is being violated by a private actor. Additionally, under the UN Guiding Principles on Business and Human Rights, companies are required to fulfil the same responsibilities for preventing, addressing, and remedying human rights abuses as state actors.

2.2.2.1 Targeted advertising

In the digital age, many of the world's largest companies make their wealth from harvesting and monetizing users' data at scale for targeted advertising.[69] Targeted advertising is a marketing strategy that delivers customised advertisements to individuals or groups online. This strategy uses tracking technologies on websites and apps to gather information about users' online behaviours – including websites visited, search queries, purchase history, how long users linger on a page, and their physical location – to predict user interests and target content that will appeal to them. Through its reliance on the collection and analysis of extensive personal data, targeted advertising raises concerns for privacy and data protection rights.

The impacts on privacy rights are intertwined with impacts on other rights, such as: non-discrimination; freedom of thought, conscience and religion; freedom of expression; right to life, liberty and security of person; and the right to free and fair elections.[70] Indeed, as a result of targeted advertising: potential job applicants have been excluded by race or gender;[71] children as young as thirteen have received harmful adverts for alcohol or weight loss pills;[72] people have been emotionally manipulated by family, friends, or co-workers;[73] people recovering from gambling addictions have received adverts promoting gambling;[74] and voters have had their political views influenced ahead of elections.[75]

In addition, targeted advertising often collects sensitive personal information without meaningful user control and consent. The right to object to personal data processing and the right to meaningfully consent to personal data processing for targeted advertising are both provided for in the UK GDPR. Specifically, for consent to be valid, it must be freely given, specific, informed, and unambiguous. For targeted advertising, consent is usually obtained through 'cookie walls' that require website visitors to click 'accept' or 'reject', or through

69 Amnesty International (2024). 'Breaking up with Big Tech: A human rights-based argument for tackling Big Tech's market power'. Amnesty International https://www.amnesty.org/en/documents/pol30/0226/2025/en/

70 Ranking Digital Rights (2019). 'Consultation Draft: Human Rights Risk Scenarios: Targetted Advertising'. Ranking DIgital Rights. https://rankingdigitalrights.org/wp-content/uploads/2019/02/Human-Rights-Risk-Scenarios-targeted-advertising.pdf

71 Propublica (2018). Facebook promises to bar advertisers from targeting ads by race or ethnicity again. Propublica https://www.propublica.org/article/facebook-promises-to-bar-advertisers-from-targeting-ads-by-race-or-ethnicity-again ; Moore (2018). How the online business model encourages prejudice. The Guardian https://www.theguardian.com/technology/2018/oct/28/how-target-ads-threaten-the-internet-giants-facebook

72 Tech Transparency Project (2021). Pill, cocktails and anorexia: Facebook allows harmful ads to target teens. https://www.techtransparencyproject.org/articles/pills-cocktails-and-anorexia-facebook-allows-harmful-ads-target-teens

73 Olson (2019). 'For $29 this man will help you manipulate your loved ones with targetted Facebook and browser links". Forbes https://www.forbes.com/sites/parmyolson/2019/01/15/a-shadowy-entrepreneur-claims-his-online-manipulation-business-is-thriving/#7adbe00972a9

74 Society for Computers & Law. (2025) 'High Court rules on targeting advertising to recovering online gambling addict'. Society for Computers & Law https://www.scl.org/high-court-rules-on-targeting-advertising-to-recovering-online-gambling-addict/

75 Saccaro (2014). "The secret experiment behind Facebook's I Voted sticker" MIC. https://www.mic.com/articles/103350/the-secret-experiment-behind-facebook-s-i-voted-sticker#.aEtJHDbgT ; Merrill (2018). 'What we learned from collecting 100,000 targeted Facebook ads'. Propublica https://www.propublica.org/article/facebook-political-ad-collector-targeted-ads-what-we-learned

'consent or pay models', but many users may not fully understand the implications of what they are consenting to when clicking 'accept'.

In 2025, two landmark legal cases challenging companies' data processing practices for targeted advertising have shown that current consent models may not meaningfully afford users their rights to consent or object to personal data processing. First, in O'Caroll vs Meta, the claimant argued that people should have the right to use Facebook without letting the company surveil or profile their personal data. O'Caroll brought the claim against Meta after she received adverts about babies on her facebook feed before she'd even told anyone that she was pregnant. Meta agreed to a settlement and afforded O'Carroll her right to turn off targeted advertising on her Facebook feed. Second, in RTM vs Bonne Terre, a recovering gambling addict brought a claim against a gambling company, arguing that he was not able to give valid consent to the advertising because of the impact of his gambling addiction. The judge's ruling in RTM vs Bonne Terre found that the claimant's decision-making was compromised by his gambling addiction, and therefore he could not properly consent. Additionally, the judge specified that context, such as the risky environment of online gambling and the subject's gambling addiction, must be taken into consideration when determining valid standards for consent to targeted advertising, and data processing more broadly.[76] This ruling is important for calling into question the standards for valid consent in relation to targeted advertising.

2.2.2.2 Generative AI

Another threat to our right to privacy by private actors is the increasing development and deployment of generative AI models. Generative AI is a form of machine learning that creates new content (eg. text, images, video, code) by learning patterns from vast amounts of existing data. One example of this is Large Language Models (LLMs) which have become common, almost everyday parts of our lives since the release of ChatGPT in November 2022. These systems are used for a wide range of applications across industries, whether healthcare, marketing and PR, education, media and entertainment, and the legal sector. Well-known commercial examples of LLMs include: OpenAI's ChatGPT, Google's Gemini, X's Grok, Anthropic's Claude, and DeepSeek.

LLMs and other generative AI models may threaten privacy rights because of the use of personal data in training; the lack of user control over data usage; and risks in relation to the absorption and regurgitation of user-inputted data.[77] LLMs are built using an enormous amount of text scraped online, from the following sources: data publicly available on the internet; data licensed from third parties; and data from users or human trainers.[78]

They raise privacy concerns because this data may involve personal information, whether from publicly available sources (eg. Wikipedia, reddit links, journals, and other sources), or from user-inputted data. One study found that 0.1% of outputted information from GPT-2 consisted of personal information such as names, addresses, or phone numbers.[79] Additionally, generative AI products often collect and process users' conversation data for the ongoing training and updating of their products – this could involve personal details such as life experiences, work status, recent thoughts, and interests.[80] This personal data from the model's training corpus may

76 AWO (2025). 'Landmark High Court ruling on GDPR consent to profiling and targeting' AWA. https://awo.agency/articles/landmark-high-court-ruling-on-gdpr-consent-to-profiling-and-targeting/?mtm_campaign=awo

77 Privacy International (2024).'Large language models and data protection' Privacy International.https://privacyinternational.org/explainer/5353/large-language-models-and-data-protection

78 OpenAI (2024). "How ChatGPT and Our Language Models are Developed: Learn More about How We Develop Our Models and Apply Them in Products like ChatGPT" Open AI. https://help.openai.com/en/articles/7842364-how-chatgpt-and-our-language-models-are-developed

79 Quach (2021). "What happens when your massive text-generating neural net starts spitting out People's phone numbers? If you're OpenAI, you create a filter" The Register https://www.theregister.com/2021/03/18/openai_gpt3_data/

80 Xiongbiao Ye, Yuhong Yan, Jia Li, Bo Jiang (2024). 'Privacy and personal data risk governance for generative artificial intelligence: A Chinese perspective',Telecommunications Policy https://www.sciencedirect.com/science/article/pii/S0308596124001484#bib31

then be leaked to other users, with profound privacy concerns.[81]

Additionally, digital rights experts have raised concerns about the lack of user control of personal data used by generative AI models. Privacy International has argued that many generative AI models do not uphold basic data subject rights from GDPR such as our right to access, rectify, or request deletion of personal data.[82] Transparency around how data is collected and processed for generative AI models is often opaque, and as such individuals are largely unaware that their personal data is being used in a model. Individuals are not informed which makes rectifying or requesting deletion difficult to near impossible.[83]

CHAPTER 2 SUMMARY POINTS:

- The right to privacy protects people against interference, both offline and online, and is therefore essential for being able to live in autonomy and dignity.
- The right to data protection expands the right to privacy by establishing comprehensive regulation for protecting people's personal data through the GDPR.
- Privacy and data protection rights are closely intertwined with other rights including non-discrimination, and freedom of expression and association.
- The digital transformation holds many concerns for the rights to privacy because many AI and data-driven technologies rely on disproportionate, unnecessary, and illegitimate data processing.
- AI and data-driven technologies by state actors that raise concerns for privacy rights include: LFR, digital ID, crime prediction systems, social media monitoring, bulk communications interceptions, and weakening encryption.
- AI and data-driven technologies by private actors that raise concerns for privacy rights include: targeted advertising and generative AI models.
- Protecting the right to privacy in the digital age therefore requires adequate enforcement of the GDPR, and robust legal frameworks governing high-risk use cases, such as digital ID and LFR. In some instances, prohibitions may be necessary on certain use cases considered incompatible with the right to privacy.

81 See Ray (2023). ChatGPT Can Leak Training Data, Violate Privacy, Says Google's DeepMind. ZDNet https://www.zdnet.com/article/chatgpt-can-leak-source-data-violate-privacy-says-googles-deepmind/ ; https://privacyinternational.org/explainer/5353/large-language-models-and-data-protection

82 Privacy International (2024). 'PI response to ICO consultation on data subject rights and generative AI' Privacy International. https://privacyinternational.org/advocacy/5338/pi-response-ico-consultation-data-subject-rights-and-generative-ai

83 Privacy International (2024). 'Privacy International's response to the Information Commissioner's Office's call for evidence on "Generative AI first call for evidence: the lawful bases for web scraping to train generative AI models"' Privacy International https://privacyinternational.org/sites/default/files/2024-03/PI%20response%20-%20ICO%20Consultation%20on%20web%20scraping%20and%20Gen%20AI%20%28submitted%29.pdf

CHAPTER 3
RIGHT TO FREEDOM OF EXPRESSION AND INFORMATION IN THE DIGITAL AGE

Article 19 of the UDHR states: *"Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers."*[84]

Article 19 of the ICCPR similarly states that, *"Everyone shall have the right to hold opinions without interference," and "to freedom of expression [which] shall include freedom to seek, receive and impart information and ideas of all kinds." The exercise of the rights provided for in paragraph 2 of this article carries with it special duties and responsibilities. It may therefore be subject to certain restrictions, but these shall only be such as are provided by law and are necessary:*

a. For respect of the rights or reputations of others;

b. For the protection of national security or of public order (ordre public), or of public health or morals."[85]

3.1 ABOUT THE RIGHT TO FREEDOM OF EXPRESSION AND INFORMATION

The rights to freedom of information and expression, as articulated in Article 19 of the UDHR and ICCPR, are crucial for living in an open, fair, and democratic society. They protect our rights to question the government and hold power accountable; attend protests; form social movements and organise politically; and communicate and connect with one another freely. Key to this is the right to access information which is embedded in freedom of expression though often overlooked. It protects individuals' ability to seek, receive, and impart essential information – which is essential for transparent governance and accountability.

84 United Nations General Assembly. (1948). Universal declaration of human rights (217 [III] A). https://www.un.org/en/about-us/universal-declaration-of-human-rights

85 United Nations General Assembly. (1966, December 16). International Covenant on Civil and Political Rights. Treaty Series, 999, 171. https://www.refworld.org/legal/agreements/unga/1966/en/17703

Like the right to privacy, the right to freedom of expression and information is not an 'absolute right', meaning it can be limited if the restriction is considered lawful, for a legitimate aim, and proportionate.

3.2 IMPACTS OF DIGITAL TECHNOLOGY ON FREEDOM OF INFORMATION, OPINION, AND EXPRESSION

This section discusses implications for freedom of expression posed by two contexts for the emergence of digital technologies: (1) online information platforms with significant influence over how people access, share, receive, and appraise information, and (2) state uses of digital surveillance technologies which can have a chilling effect on free expression if people are driven to modify their behaviour out of fear – such as by remaining silent, avoiding certain places, or eschewing protests and demonstrations which are critical avenues for preserving democracy. We will discuss each in turn.

3.2.1. Online platforms and restricted access to information:

The emergence of online platforms has significantly impacted how we access information in the digital age, massively expanding possibilities for accessing information online. In particular, online social media platforms such as Instagram, Facebook, X, Reddit, and Discord have become key sites where people access valuable information, such as news and political information, advice on abuse, abortion, or mental and physical health issues. This is especially so for young people; research published by Ofcom in July 2025 found that 80% of 16-24-year-olds went online to get their news, with 75% looking specifically to social media.[86]

However, the design of online platforms, and the technical and regulatory systems and procedures in place that determine the flow of information online, pose concerns regarding the human right to receive and access information without interference. Free expression advocates from the digital rights community have been raising concerns about the potential impacts of online platforms' content moderation systems on our access to important information.[87] Indeed, there are documented examples where lawful information on online platforms has been unnecessarily removed or demoted (known as 'shadow-banning'), thereby potentially threatening to undermine our right to access information that may be important for the democratic health of society, and for people's mental and physical wellbeing. With platforms effectively acting as online public spaces in the digital age, the concerns for the right to free expression and access to information are notable.

The unnecessary removal of information may take place because the content moderation systems on online platforms that are intended to limit the circulation of harmful content online – such as abuse, hate speech, or suicide content – take an overly lenient approach. While effective and well-resourced content moderation systems are vital for ensuring that information online does not infringe on our rights – such as non-discrimination, freedom from inhuman or degrading treatment, or children's right to protection from sexual abuse and exploitation – the flipside is that overly lenient approaches may result in limiting access to vital information. Indeed, some examples highlighted by the digital rights community where access to high-quality and essential information such as health information include the removal of: content

86 Ofcom (2025, July 21). Top trends from our latest look at the UK's news habits. https://www.ofcom.org.uk/media-use-and-attitudes/attitudes-to-news/top-trends-from-our-latest-look-at-the-uks-news-habits

87 Amnesty International. (2024, June 19). United States: Social media companies' removal of abortion-related content may hinder access to accurate health information. https://www.amnesty.org/en/latest/news/2024/06/united-states-social-media-companies-removal-of-abortion-related-content-may-hinder-access-to-accurate-health-information/

about the Israeli occupation and violence in Palestine;[88] abortion-related content;[89] as well as lawful sexually-themed content, especially of LGBTQ+ communities.[90]

Restricted access to important information on online platforms can take place due to both human and automated content moderation systems. Indeed, platforms are increasingly replacing human and professional content moderation teams with AI systems for identifying and removing potentially unlawful or harmful content.[91] However, evaluative research on AI-moderation systems has found poor ratings in terms of their accuracy, consistency, and bias which may contribute to the removal of lawful and proportionate information online.[92] Meanwhile, human moderation teams may over-moderate lawful content because some rules on content categories may be difficult to enforce consistently – especially when it comes to categories of content that may be difficult to define or establish concrete parameters around, such as 'hate speech' or 'misinformation', which also often require high degrees of specific understandings of local, geographic, linguistic, or contextual knowledge.[93]

State authorities may also play a role in the moderation of online content. Government bodies such as law enforcement authorities sometimes hold 'trusted flagger' status with platforms to identify illegal content for expedited content take-downs. For example, London's Metropolitan Police has trusted flagger status with YouTube,[94] and the Home Office's Counter Disinformation Unit (CDU) which has since been disbanded held trusted flagger status for monitoring Covid-19 disinformation.[95] In both these examples, free expression advocates among the digital rights community raised concerns about unnecessary and disproportionate state involvement in the removal of lawful content. In 2021, the Met referred 510 music videos on YouTube for removal, mainly rap and drill videos, for incitement of violence, which the Electronic Freedom Foundation stated was an unnecessary infringement on people's right to creative expression, disproportionately affecting young Black men.[96] Similarly, Big Brother Watch found that only 42 per cent of the CDU's content reports to X/Twitter actually breached the platform's Terms of Service, citing an overreach of state involvement in platform moderation.[97]

Overall, content moderation on online platforms raises complex procedural challenges for platforms and regulatory initiatives in balancing people's rights. Effective content moderation systems are essential for protecting an array of human rights online, and failures to moderate discriminatory content such as hate speech or abuse can lead to online and offline harm – such as X's Community Notes system during the 2024 racist riots which broke out in Southport, as discussed in Chapter 1.[98] However, ineffective or disproportionate content moderation could

88 Electronic Frontier Foundation (2025, April 25). Platforms must stop unjustified takedowns of posts by and about Palestinians. https://www.eff.org/deeplinks/2023/11/platforms-must-stop-unjustified-takedowns-posts-and-about-palestinians ; Younes, R. (2024). Meta's broken promises. Human Rights Watch. https://www.hrw.org/report/2023/12/21/metas-broken-promises/systemic-censorship-palestine-content

89 Electronic Frontier Foundation. (2025, September 13). Our stop censoring abortion campaign uncovers a social media censorship crisis. https://www.eff.org/pages/our-stop-censoring-abortion-campaign-uncovers-social-media-censorship-crisis ; Amnesty International. (2024, June 19). United States: Social media companies' removal of abortion-related content may hinder access to accurate health information. https://www.amnesty.org/en/latest/news/2024/06/united-states-social-media-companies-removal-of-abortion-related-content-may-hinder-access-to-accurate-health-information/

90 Electronic Frontier Foundation (2025, February 5). Meta's new content policy will harm vulnerable users. https://www.eff.org/deeplinks/2025/01/metas-new-content-policy-will-harm-vulnerable-users-if-it-really-valued-free

91 Kerr, D. (2025, August 10). TikTok to replace trust and safety team in Germany with AI and outsourced labor. The Guardian. https://www.theguardian.com/technology/2025/aug/10/tiktok-trust-safety-team-moderators-ai

92 Boucher, H. (2025, September 15). AI models are struggling to identify hate speech, study finds. The Independent. https://www.independent.co.uk/news/uk/home-news/ai-hate-speech-study-university-pennsylvania-b2826860.html

93 York, J. (2020, April 7). The global impact of content moderation. Article 19. https://www.article19.org/resources/the-global-impact-of-content-moderation/ ; The Santa Clara Principles On Transparency and Accountability in Content Moderation. https://santaclaraprinciples.org/

94 In 2018, the Met Police became the first law enforcement authority globally to receive trusted flagger status. See: Pritchard, W., & Pritchard, W. (2024, July 27). YouTube is Working With Met Police to Take Down Rap and Drill Videos. VICE. https://www.vice.com/en/article/met-police-youtube-drill-music-removal/

95 Big Brother Watch. (2023). Fact Checking the Government's 'Fact Sheet' on the Counter Disinformation Unit. https://bigbrotherwatch.org.uk/blog/fact-checking-govt-fact-sheet/

96 Collings, P. (2022, October 5). How YouTube's Partnership with London's Police Force is Censoring UK. Electronic Frontier Foundation. https://www.eff.org/deeplinks/2022/08/how-youtubes-partnership-londons-police-force-censoring-uks-drill-music

97 Big Brother Watch. (2023). Fact Checking the Government's 'Fact Sheet' on the Counter Disinformation Unit. https://bigbrotherwatch.org.uk/blog/fact-checking-govt-fact-sheet/

98 Perry, H. and Malik, N. (2025). Researching the riots: An evaluation of the efficacy of Community Notes during the 2024 Southport riots. Demos. https://demos.co.uk/research/researching-the-riots-an-evaluation-of-the-efficacy-of-community-notes-during-the-2024-southport-riots/

also undermine people's rights to access information by unduly removing lawful information. As such, content moderation systems require a careful balancing of people's rights in both platform policies and procedures, and government legislation for the regulation of online speech.[99]

3.2.2. Chilling effects of surveillance technologies on free expression:

This section will now discuss the increasing usage of AI, data-driven and digital technologies for surveillance purposes and their potential 'chilling effects' on free expression. A chilling effect occurs when people refrain from exercising their rights or engaging in lawful behavior because they fear negative consequences.

As identified in the first chapter on privacy, state authorities are increasingly incorporating AI, algorithmic, and data-driven tools into their activities. These include a growing number of surveillance technologies such as crime prediction systems,[100] social media monitoring,[101] and Live Facial Recognition (LFR). In July 2025, the Met announced that it was doubling its usage of LFR up to 10 van deployments every week,[102] and fixed location LFR is also becoming increasingly common.[103] Already in 2024, 4.5million faces across the UK were scanned by LFR.[104]

This roll-out is happening even though there is awareness of the potential implications for people's freedom of expression. Indeed, a 2019 audit of LFR commissioned by the Met police conducted by Dr Darragh Murray states:

> *"Importantly, the deployment of LFR technology may generate a chilling effect, whereby individuals refrain from lawfully exercising their democratic rights due to a fear of the consequences that may follow. This may harm a number of rights, including the right to freedom of expression, the right to freedom of assembly and association, and the right to freedom of religion."*[105]

Evidence gathered by a police oversight body also demonstrates that people may not feel safe to openly speak, communicate, or interact with others in public in locations where LFR is being deployed. A London Policing Ethics Panel report on LFR found that 38% of 16-24 year-olds would stay away from events or places where facial recognition surveillance was being used, as well as high numbers of Black, Asian and Minority Ethnic people.[106] This may be because people are concerned by the privacy implications resulting from the indiscriminate collection of their sensitive biometric information in public spaces, and the possibility of algorithmic bias and discrimination, whereby there is a higher likelihood of people of colour being wrongly flagged. The potential consequences of being flagged by LFR are serious – and has resulted in people

99 The Santa Clara Principles, drafted by a group of digital rights organisations in 2018, offer recommendations to better ensure that the enforcement of content guidelines is fair, unbiased, proportional, and respectful of users' rights, including on: human rights and due process; understandable rules and policies; cultural, linguistic, and contextual understanding; state involvement in content moderation; integrity and explainability; access to data; notices; and appeals. See: https://santaclaraprinciples.org/

100 Amnesty International UK. (2025). AUTOMATED RACISM: How police data and algorithms code discrimination into policing. https://www.amnesty.org.uk/files/2025-02/Automated%20Racism%20Report%20-%20Amnesty%20International%20UK%20-%202025.pdf

101 Privacy International. (2024). Social media monitoring in the UK: the invisible surveillance tool increasingly deployed by government. https://privacyinternational.org/long-read/5337/social-media-monitoring-uk-invisible-surveillance-tool-increasingly-deployed

102 Badshah, N. (2025, August 8). Met police to more than double use of live facial recognition. The Guardian. https://www.theguardian.com/technology/2025/jul/31/met-police-to-more-than-double-use-of-live-facial-recognition

103 In Cardiff, South Wales police have implemented a ring of 11 fixed LFR cameras across the city for switching on at busy periods, such as during a Beyonce concert in 2023 or the Six Nations Rugby tournament in 2025. The Met is also planning to install fixed LFR in Croydon, London. See: https://www.south-wales.police.uk/news/south-wales/news/2025/february/extra-live-facial-recognition-cameras-cardiff-city-centre-keep-visitors-safe/ ; Galliven, B. (2025, July 12). Croydon's fixed facial recognition cameras spark debate. BBC News. https://www.bbc.co.uk/news/articles/cy0w5egz91no

104 France24. (2025, August 24). UK's mass facial-recognition roll-out alarms rights groups. https://www.france24.com/en/live-news/20250824-uk-s-mass-facial-recognition-roll-out-alarms-rights-groups

105 Fussey, P., Murray, D., & University of Essex. (2023). Regulating biometrics: global approaches and urgent questions. In Regulating Biometrics: Global Approaches and Urgent Questions (pp. 78–80). https://ainowinstitute.org/wp-content/uploads/2023/09/regulatingbiometrics-fussey-murray.pdf

106 Big Brother Watch (2020). Briefing on facial recognition surveillance. https://bigbrotherwatch.org.uk/wp-content/uploads/2020/06/Big-Brother-Watch-briefing-on-Facial-recognition-surveillance-June-2020.pdf

being wrongfully stopped, interrogated, searched, and in some cases arrested.[107]

Because of the concerns LFR presents to free expression – as well as privacy (Chapter 2), non-discrimination (Chapter 1) – many digital rights groups and international bodies are calling for the prohibition of LFR in public spaces (See Chapter 2 Table 1) and in 2024, the UN Human Rights Committee concluded that the UK government should "end the use of facial recognition and other mass surveillance technologies by law enforcement agencies at protests, to safeguard privacy, non-discrimination, freedom of expression, association, and assembly rights for protestors".[108]

Many of the other surveillance technologies used by the UK government raise similar concerns for free expression as LFR. For example, on the use of geographic crime prediction technologies by UK police forces, Liberty has stated:

> *"As we normalise predictive policing, we may begin to self-police to avoid unwarranted suspicion. We may become afraid of the level of data being gathered about us, what it is used for, how it is shared and what predictions might be made about us as a result – and this may have a chilling effect on what we choose to say, where we choose to go and who we choose to associate with."*[109]

CHAPTER 3 SUMMARY POINTS:

- The right to freedom of expression protects people's ability to hold and express their own opinions, and as such is vital in a healthy democracy. It also protects people's ability to access and receive information.
- In the digital age, online information ecosystems may impact our right to access information because ineffective or overly lenient content moderation systems may remove lawful information.
- Surveillance technologies such as LFR or crime prediction systems may result in chilling effects on free expression if people fear the consequences and outcomes of state surveillance.
- Protecting the right to freedom of expression and information in the digital age therefore requires effective and well-resourced content moderation systems to ensure discriminatory or harmful content is removed without over-moderating lawful content. It may also require robust legal frameworks governing surveillance technologies and, in some instances, prohibitions on certain use cases considered incompatible with the right to freedom of expression.

107 Shaun Thompson, an anti-knife crime campaigner, is currently bringing a High Court challenge against the Met after he was wrongfully identified as a suspect. See: Jessup, S. (2025, August 6). "Met Police facial recognition tech mistook me for wanted man." BBC News. https://www.bbc.co.uk/news/articles/cqxg8v74d8jo

108 Newson, N. (2024, April 18). UN standards on the use of surveillance technology at protests. House of Lords Library. https://lordslibrary.parliament.uk/un-standards-on-the-use-of-surveillance-technology-at-protests/

109 Liberty. (2020, March 4). Report: Policing by machine. https://www.libertyhumanrights.org.uk/issue/policing-by-machine/

CHAPTER 4
RIGHT TO AN EFFECTIVE REMEDY IN THE DIGITAL AGE

Article 8 of the UDHR states: *"Everyone has the right to an effective remedy by the competent national tribunals for acts violating the fundamental rights granted him by the constitution or by law."*[110]

Article 2 (a-c) of the ICCPR similarly mandates that *"any person whose rights or freedoms… are violated shall have an effective remedy… determined by competent judicial, administrative or legislative authorities,"* and that *"competent authorities shall enforce such remedies when granted."*[111]

4.1 ABOUT THE RIGHT TO AN EFFECTIVE REMEDY

The right to an effective remedy gives individuals the free-standing fundamental right to seek legal remedies after their human rights have been violated. For this reason, it plays a crucial role in operationalising all other human rights.

For a remedy to be 'effective', the UN Committee on Economic, Social and Cultural Rights (CESCR) has highlighted that they must be 'accessible', 'affordable', 'timely', and 'effective'.[112] Examples where the European Court of Human Rights (ECtHR) has found the UK to be in breach of the right to an effective remedy include: *Chalkley v. The United Kingdom (2003)* where the applicant found there was no remedy available at a national level when he complained about the privacy violation of police installing a covert surveillance device in his home;[113] *Armstrong v. The United Kingdom (2002)* when police undertook covert audio surveillance without a legal basis, meaning that the applicant lacked an effective domestic remedy for the breach of his privacy rights;[114] and *M.A.K. & R.K. v. The United Kingdom (2010)* when the withdrawal of the applicant's legal aid deprived her of an effective remedy following hospital treatment that she argued took place without consent.[115]

110 United Nations General Assembly. (1948). Universal declaration of human rights (217 [III] A). https://www.un.org/en/about-us/universal-declaration-of-human-rights
111 United Nations General Assembly. (1966, December 16). International Covenant on Civil and Political Rights. Treaty Series, 999, 171. https://www.refworld.org/legal/agreements/unga/1966/en/17703
112 See UN Committee on Economic, Social and Cultural Rights, General Comment No. 9: The domestic application of the Covenant, 3 December 1998, paragraph 9
113 Chalkley v. The United Kingdom, App. No. 63831/00 (Eur. Ct. H.R. Sept. 26, 2002). https://hudoc.echr.coe.int/eng?i=001-22695
114 Armstrong v. The United Kingdom, 65282/09 ECHR (2014). https://hudoc.echr.coe.int/eng?i=001-148670 HUDOC - European Court of Human Rights
115 M.A.K. and R.K. v. The United Kingdom, 45901/05 and 40146/06 (ECHR Mar. 23, 2010). https://hudoc.echr.coe.int/eng?i=001-97880

This chapter explores the potential impacts of new and emerging digital technologies on the right to effective remedy. So far, this report has outlined ways that technology and AI hold threats for our fundamental human rights. In the digital age, the right to an effective remedy ensures we are able to seek effective redress from rights infringements that are caused or connected to technology. There are many related and interconnected concepts related to effective remedies that are drawn on by the digital rights community including: redress, accountability, recourse, and contestability. The terms can largely be used interchangeably.

4.2 IMPACT OF DIGITAL TECHNOLOGY ON EFFECTIVE REMEDIES

New technologies such as AI are creating various challenges for people seeking effective remedies for tech-induced harms, such as those outlined so far in this report. The challenges drawing significant attention among the digital rights community are: transparency issues and liability challenges stemming from complex digital supply chains. These challenges present urgent questions, upon which the operationalisation of all remedies for tech harms depends. To summarise some of these challenges, this section draws on the work on redress of Dr. Yulu Pi, an AI Governance researcher at the Research Center Trustworthy Data Science and Security, who focuses on explainability within regulatory frameworks and from technical and design perspectives.[116]

4.2.2 Transparency

A crucial part of obtaining an effective remedy relies on having access to transparent information around a technology system, including how it works, is developed and deployed. Transparency is necessary so that people impacted by digital rights abuses are able to identify and evidence the cause and pinpoint responsibility for the potential rights infringements.

However, examples in this report have shown that tech systems, and especially AI systems, are often designed and deployed without sufficient transparency for people to easily seek redress. For example, people have struggled to access information about automated decision-making systems used in the public sector (such as the DWP's fraud prediction tools or police forces' crime prediction systems), recommender algorithms on social media, and LLM chatbots. Without this rudimentary first step of transparency, redress is not possible.

Transparency issues related to technology can occur on a number of levels and by a range of actors, all of which present barriers for individuals seeking redress.

Technical opacity we use here to refer to the incomprehensibility of a technological system to human users. It is a particular concern with respect to AI systems where even the expert developers may not fully understand the intricacies of how an AI system functions. With machine learning algorithms, the decision-making process is often so complex that researchers may be able to explain the data inputs and outputs, but struggle to precisely explain the inner workings of the model.[117] Technical opacity poses a challenge for establishing causality for tech harms.

Capability opacity we use to refer to the difficulty of acquiring sufficient information about a digital technology's functional capacity to be able to seek effective remedies after potential rights infringements have occurred. This information may include: a system's intended purposes, how it operates in practice, the data types and sources used, what decisions or outcomes it influences, any performance metrics, reliability, or limitations.

116 Pi, Y., & Proctor, M. (2025). Toward empowering AI governance with redress mechanisms. Cambridge Forum on AI: Law and Governance. https://www.semanticscholar.org/paper/Toward-empowering-AI-governance-with-redress-Pi-Proctor/dbd5cf4d6126b58af6af4e43d919fc1915509a6b

117 Savage, N. (2022). Breaking into the black box of artificial intelligence. Nature. https://doi.org/10.1038/d41586-022-00858-1

Organisational opacity we use here to refer to institutional practices, processes, or structures making it difficult for outsiders (and sometimes insiders) to see and understand information about the organisation's behaviors. In his book *Black Box Society*, Frank Pasquale describes how a number of organisations actively withhold information about their organisational practices using AI in particular. In some instances, decisions not to provide certain pieces of information may be justified in the interests of protecting IP or for example retaining anonymity of individual data, however, opaque informational practices pose significant barriers for rights-holders to seek redress and accountability where they may have had their rights infringed by technology usage. Indeed, we have seen instances of organisational opacity in a number of examples in this report. In relation to state actors, a key example is the challenges faced by civil society organisations when undertaking Freedom of Information Act (FOIA) requests. For Amnesty UK's 'Automated Racism' report on crime prediction algorithms used by UK police forces, information about the algorithms in use was obtained through multiple rounds of FOIAs to all 43 police forces in the country which is a long, arduous, and costly process, with responses often refusing to provide information.[118]

In relation to private actors, a key example is the difficulties in obtaining information about recommender algorithms used by online platforms such as X – evaluative research on the impacts of recommender algorithms, such as by the ISD or Molly Rose Foundation usually rely on the creation of dummy accounts because of challenges accessing key information needed for redress such as training datasets or model parameters.[119]

Finally, it is important to note that opacity may result in an individual being unaware that the infringement on their rights is due to a technology system in the first place. In many cases, this causality is not apparent: people often do not know that they have been harmed by a tech system, used both by corporate and state actors. For example, people who were flagged for investigation by the DWP's fraud detection algorithm initially were unaware that they were being investigated because of an outcome from a risk-scoring algorithm.[120] Similarly, increased people may be unaware that increased policing in their local area is because the area has been flagged as a 'crime hot spot' by a geographic crime prediction algorithm.[121]

4.2.2.1 Overcoming transparency challenges

Given these challenges that technical and organisational opacity bring for people's ability to seek effective remedies, the digital rights community are calling for regulation that requires transparency from both public sector bodies and private companies' uses of technology.[122]

At the level of national policy, many governments have responded by introducing mandatory transparency measures. For example, the EU AI Act includes transparency obligations on companies requiring providers of 'high-risk' AI systems to explain to deployers how the system works, and how data is processed, and providers of certain general purpose AI systems must inform users that they are interacting with an AI system.[123] California and New York have

118 Amnesty International UK. (2025). AUTOMATED RACISM: How police data and algorithms code discrimination into policing. https://www.amnesty.org.uk/files/2025-02/Automated%20Racism%20Report%20-%20Amnesty%20International%20UK%20-%202025.pdf

119 Institute of Strategic Dialogue. (2025, July 16). Towards transparent recommender systems: Lessons from TikTok research ahead of the 2025 German federal election. https://www.isdglobal.org/digital_dispatches/towards-transparent-recommender-systems-lessons-from-tiktok-research-ahead-of-the-2025-german-federal-election/ ; Molly Rose Foundation. (2025, January 22). New research exposes tech giants' amplification of content promoting suicide and self-harm. https://mollyrosefoundation.org/new-research-exposes-tech-giants-amplification-of-content-promoting-suicide-and-self-harm/

120 Hegarty, T. (2024, June 17). NEW CASE: secret algorithm targets disabled people unfairly for benefit probes – cutting off life-saving cash and trapping them in call centre hell. Foxglove. https://www.foxglove.org.uk/2021/12/01/secret-dwp-algorithm/

121 Amnesty International UK. (2025). AUTOMATED RACISM: How police data and algorithms code discrimination into policing. https://www.amnesty.org.uk/files/2025-02/Automated%20Racism%20Report%20-%20Amnesty%20International%20UK%20-%202025.pdf

122 Amnesty International UK. (2025). AUTOMATED RACISM: How police data and algorithms code discrimination into policing. https://www.amnesty.org.uk/files/2025-02/Automated%20Racism%20Report%20-%20Amnesty%20International%20UK%20-%202025.pdf ; Public Law Project. (2024). Securing meaningful transparency of public sector AI. https://publiclawproject.org.uk/content/uploads/2024/10/Securing-meaningful-transparency-of-public-sector-AI.pdf

123 Key issue 5: Transparency obligations. EU AI Act. https://www.euaiact.com/key-issue/5

introduced frontier AI transparency obligations through the SB53 and RAISE Acts requiring companies to publish their safety policies, which must include policies on how to test for and respond to certain kinds of AI risks.[124]

Amnesty International adds that government bodies should also be obliged to publish the following information: intended purpose of their system; how the system operates in practice; all data types and data sources used by the system; what decisions or outcomes the system influences; and any internal reviews or evaluations.[125]

The UK government has also made an attempt to improve transparency around its use of AI. Government departments are required to upload information about AI systems they are using onto the Algorithmic Recording Transparency Standard (ATRS). While this is a good start, the actual level of information provided is limited. As of October 2025, there were only 89 records uploaded across the whole UK government.The actual number is likely far higher – indeed one report found from 2018-2025, the UK public sector awarded 1,309 AI contracts.[126] This suggests that many government departments are not reporting their AI systems on the ATRS.

Public Law Project see the potential value in ATRS and have undertaken valuable work on transparency around public sector uses of AI in the UK. They recommend putting the ATRS on a statutory footing to strengthen legal obligations for public sector authorities to disclose what systems they are using, and that individuals should be notified when a public sector authority uses an AI, algorithmic, or automated tool in a decision about them.[127] In addition, there should not be transparency exemptions for law enforcement, immigration, national security or counterterrorism bodies.

These transparency measures would be a vital starting point for enabling people to seek effective remedies from tech-induced harms.

4.2.3 Complex digital value chains: The "many hands" problem

Transparency is only the first hurdle for people seeking redress. Even if people are able to identify that they have been harmed and access information about the mechanism of harm, the next challenge is identifying who is responsible for the harm and for providing remedy.

Interviewees for this project characterised the primary difficulty of pinpointing responsibility for harms caused by more complex digital technologies like AI as the "many hands" problem. The challenge is arising from intricacies of the AI value chain. This problem is summarised by the European Parliament's former EU AI Liability Directive: "The large number of people potentially involved in the design, development, deployment, and operation of high-risk AI systems, makes it very difficult for plaintiffs to identify the person potentially liable for damage caused and to prove the conditions for a claim for damages."[128]

Indeed, a violation of a collective or individual's rights could arise at multiple stages of the AI supply chain, whether the model developer neglecting to adequately safety test its models, a service provider who failed to adequately safety test its models, a deployer who neglected to

124 Gluck, J. (2026, January 8). The RAISE Act vs SB 53: A Tale of Two Frontier AI Laws. Future of Privacy Forum. https://fpf.org/blog/the-raise-act-vs-sb-53-a-tale-of-two-frontier-ai-laws/

125 Amnesty International UK. (2025). AUTOMATED RACISM: How police data and algorithms code discrimination into policing. https://www.amnesty.org.uk/files/2025-02/Automated%20Racism%20Report%20-%20Amnesty%20International%20UK%20-%202025.pdf

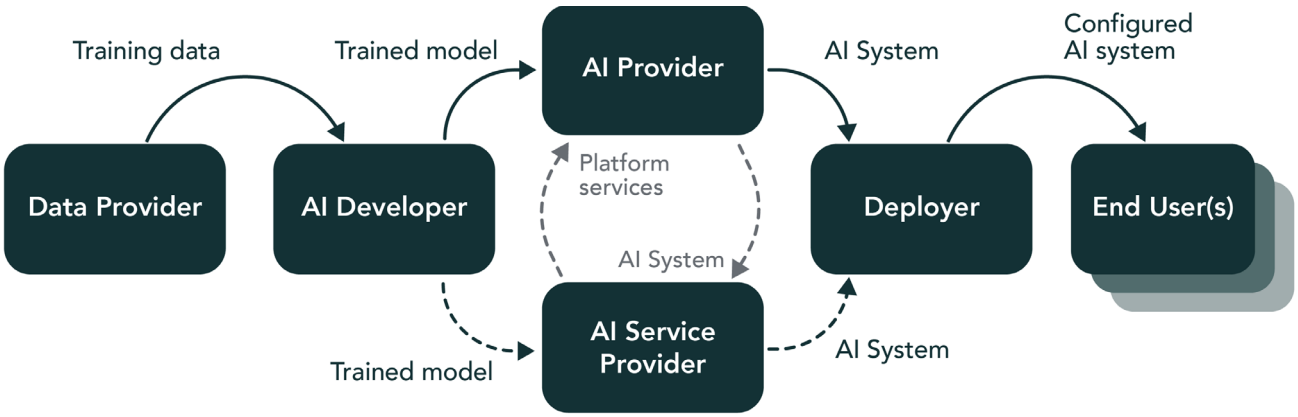
126 Horton, C. (2025, September 4). UK public sector's £3.45 billion AI spend – but it's still only a fraction of all IT contracts. THINK Digital Partners. https://www.thinkdigitalpartners.com/news/2025/09/04/uk-public-sectors-3-45-billion-ai-spend-but-its-still-only-a-fraction-of-all-it-contracts/

127 Public Law Project. (2024). Securing meaningful transparency of public sector AI. https://publiclawproject.org.uk/content/uploads/2024/10/Securing-meaningful-transparency-of-public-sector-AI.pdf

128 European Parliament. (2023). EU Artificial intelligence liability directive. https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/739342/EPRS_BRI(2023)739342_EN.pdf

undertake adequate impact assessments for its use case, or an end user misusing the product, or some combination thereof. Figure 1 illustrates the many actors comprising the AI value chain.

FIGURE 1
THE AI VALUE CHAIN[129]



As Dr Yulu Pi writes: "This dynamic enables developers and deployers to evade responsibility, leaving the human user to bear the brunt of the consequences. [...] The fragmentation of control and responsibility across those sectors in the AI value chain leads to a "many hands problem," where no one is responsible for outcomes which multiple people helped produce." [130]As a consequence, individuals and communities seeking redress from potential tech-induced rights infringements may struggle to identify who is liable, and therefore which pathways to redress are available to them.

CHAPTER 4 SUMMARY POINTS:

- The right to an effective remedy gives individuals the free-standing fundamental right to seek redress where their rights have been violated.
- In the digital age, transparency issues of AI and data-driven technologies create challenges for people to seek redress after potential rights infringements.
  - Transparency issues can prevent individuals from knowing that harm has been caused by a tech system.
  - Technical opacity means that even experts may not understand how an AI or machine learning system works.
  - Capability opacity means that people may struggle to access sufficient information about a tech system that is required to seek redress.

129 AI Value Chain — International AI Governance Association. https://intaigovassoc.org/ai-value-chain
130 Pi, Y., & Proctor, M. (2025). Toward empowering AI governance with redress mechanisms. Cambridge Forum on AI: Law and Governance. https://www.semanticscholar.org/paper/Toward-empowering-AI-governance-with-redress-Pi-Proctor/dbd5cf4d6126b58af6af4e43d919fc1915509a6b

    - Organisational opacity means that institutional practices and systems may create barriers for outsiders (and insiders) to understand how a system works.
- Subsequently assigning responsibility for tech harms is further complicated by the "many hands problem": the complexity of the AI value chain in particular makes it difficult to determine who along that chain should be held liable for any harms or rights infringements caused by the technology.
- Protecting the human right to effective remedy in our digital age will require adequate transparency around the development, use, and governance of digital technologies. It will also require the development of effective distributed liability legal frameworks that can accommodate the complexities of digital value chains.

CHAPTER 5
RIGHT TO SOCIAL SECURITY IN THE DIGITAL AGE

The right to social security is articulated in Article 22 of the UDHR: *"Everyone, as a member of society, has the right to social security and is entitled to realization, through national effort and international co-operation and in accordance with the organization and resources of each State, of the economic, social and cultural rights indispensable for his dignity and the free development of his personality."*[131]

5.1 ABOUT THE RIGHT TO SOCIAL SECURITY

The right to social security obliges states to do everything they can to respect, protect and fulfill the basic needs of citizens when they are facing life's challenges. It is deeply connected to the right to an adequate standard of living and to the idea of the welfare state. States must ensure that citizens can access benefits when they are unable to work (whether due to old age, unemployment, sickness, or caring for dependencies) and that everybody can access health services.[132] The right to social security requires that everybody should be able to access social welfare without discrimination, "especially individuals belonging to the most disadvantaged and marginalized groups".[133]

The right to social security was initially adopted in 1948 in the Universal Declaration, and has since then been turned into a legally binding obligation in the International Covenant on Economic, Social and Cultural Rights, ratified by the UK 1976.[134] Social security is mentioned in at least 119 constitutions around the world.

At the time of the Universal Declaration, there was enormous political commitment in the UK to realizing the goals of social security. Within a few years, the post-war Labour government had introduced a number of extensive social reforms to ensure universal access to healthcare and education, and a comprehensive social security system including pensions and unemployment support.

131 United Nations General Assembly. (1948). Universal declaration of human rights (217 [III] A). https://www.un.org/en/about-us/universal-declaration-of-human-rights
132 Committee on Economic, Social and Cultural Rights. (2008). General comment No. 19: The right to social security (Article 9). United Nations. https://www.ohchr.org/en/documents/general-comments-and-recommendations/general-comment-no-19-right-social ; https://www.refworld.org/legal/general/cescr/2008/en/41968
133 Ibid.
134 United Nations General Assembly. (1966). International Covenant on Economic, Social and Cultural Rights. United Nations. https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-economic-social-and-cultural-rights

5.2 IMPACTS OF DIGITAL TECHNOLOGY ON SOCIAL SECURITY

The years that have followed have brought many changes to the UK's social security system with the arrival of the digital age bringing specific and new challenges. The Department for Work and Pensions, as part of a wider government drive for technological innovation in the public sector,[135] is making significant investments to digitise the welfare system.[136] Indeed, the DWP's technology spending in 2023-24 was roughly £1 billion.[137] Examples of DWP technology procurement include: a conversational AI helpline;[138] cloud computing programmes for service delivery;[139] predictive analytics tools for fraud detection; automated payment systems; data platforms and more.[140] In the 2025 budget, the Labour government also announced a crackdown on benefit fraud which is likely to include an increase in fraud detection algorithms.[141]

Some of these uses of technology within the welfare system are having a positive impact on some people's right to social security, for example, by improving delivery and accessibility of services. However, digital rights organisations are expressing concerns that other uses are unfairly withholding people's access to social security because of digital exclusion and discriminatory automated decision-making systems.[142] As such, the increasing digitisation of the UK's welfare system calls for a re-evaluation of how to ensure the fair and nondiscriminatory enjoyment of the human right to social security in our digital era.

Drawing on recent research and interviews from digital rights experts, this section evaluates some of the key impacts of the "digital welfare state" for the right to social security through two use cases of digital technology by the DWP: the 'digital by default' benefit systems such as Universal Credit; and a range of automated decision-making algorithms used to detect fraud.[143]

5.2.1 'Digital by default' benefit systems

When Universal Credit was introduced in 2012, it was set up as the first 'digital by default' benefit system. 'Digital by default' means that digital mechanisms are now the primary mode for benefit claimants to manage application and payment processes.[144] The 'digital by default' Universal Credit system was introduced to replace in-person 'signing-on' processes in job centres. It requires claimants to create an online account, fill out an online application form, and carry out regular online tasks, such as communicating with their work coach, reporting changes in circumstances and job search updates, and monitoring payments. To set up their account,

135 United Kingdom Government. (2025). 'AI opportunities action plan' Department for Science, Innovation and Technology https://www.gov.uk/government/publications/ai-opportunities-action-plan/ai-opportunities-action-plan

136 United Kingdom Government. (2024). 'Executive Summary: Digitalising Welfare Services' Department for Work and Pensions. https://www.gov.uk/government/publications/review-of-international-and-private-sector-evidence-on-the-effectiveness-of-digitising-services/executive-summary-digitalising-welfare-services

137 Trendall, S. (2024). 'DWP increases annual digital spending to £1bn after 8% rise in FY24'. Public Technology. https://www.publictechnology.net/2024/08/14/society-and-welfare/dwp-increases-annual-digital-spending-to-1bn-after-8-rise-in-fy24/

138 Trendall, S. (2025). 'DWP plans £11m engagement to ensure long-term future of AI-powered 'conversational platform'. Public Technology https://www.publictechnology.net/2025/06/27/society-and-welfare/dwp-plans-11m-engagement-to-ensure-long-term-future-of-ai-powered-conversational-platform/

139 DWP Digital. (no date). 'Enhancing our software engineering capabilities with cloud technology'. Department of Work and Pensions Digital. https://careers.dwp.gov.uk/our-teams/dwp-digital-engineering/software-engineering-cloud-technology/

140 McCully, J. (2020). 'Explainer: What is the digital welfare state.' Digital Freedom Fund. https://digitalfreedomfund.org/explainer-what-is-the-digital-welfare-state/

141 Wright, J. (2025). 'Major DWP benefit fraud crackdown announced in budget.' The National. https://www.thenational.scot/news/national/uk-today/25651648.major-dwp-benefit-fraud-crackdown-announced-budget/

142 Amnesty International. (2025). 'Too much technology not enough empathy': How the UK's push to digitise social security harms human rights'. Amnesty International. https://www.amnesty.org/en/documents/eur45/9478/2025/en/ ; Vincent, R. (2025). 'Suspicion by design'. Big Brother Watch, https://bigbrotherwatch.org.uk/wp-content/uploads/2025/07/Suspicion-By-Design-2.pdf ; Guo, E. Geiger, G., Braum, JC. (2025). 'Inside Amsterdam's high-stakes experiment to create fair welfare AI'. MIT Technology Review https://www.technologyreview.com/2025/06/11/1118233/amsterdam-fair-welfare-ai-discriminatory-algorithms-failure/

143 Special Rapporteur on extreme poverty and human rights. (2019, October 11). A/74/493: Digital welfare states and human rights — Report of the Special Rapporteur on extreme poverty and human rights. Office of the United Nations High Commissioner for Human Rights. https://www.ohchr.org/en/documents/thematic-reports/a74493-digital-welfare-states-and-human-rights-report-special-rapporteur

144 Parliament UK. (2021). 'Welfare Policy in Scotland' Parliament UK https://publications.parliament.uk/pa/cm5802/cmselect/cmscotaf/55/5508.htm

claimants need internet access, an internet-enabled device, an email address, and a mobile phone number to receive verification codes.[145]

While the 'digital by default' benefit system beneficially facilitates the right to social security for some people by processing claims more quickly and efficiently, for others, it has the potential to undermine the right to social security where it is not accessible to groups who are digitally excluded. In the UK, a large portion of the population do not have access to the internet, predominantly because of affordability of internet devices, such as a computer or smart phone, or connectivity through broadband or phone data. In 2018, the Office for National Statistics found that 10% of the UK adult population were non-internet users,[146] while research by Ofcom in 2021 found that approximately 2 million UK households do not have home internet access[147] which leaves people without a way to access online services from their place of residence, including 'digital by default' public services. People might be able to visit a public library to access public services online, but as Amnesty International writes, " [public libraries] are not accessible to all. Access may be particularly difficult for people with a disability, people who live in rural areas, for whom the cost of travelling to a library may be prohibitive, and people whose local public library has closed."[148]

Even if access to internet and digital devices is acquired, 'Digital by default' systems also require users to wield a certain level of digital skills to make effective use of the online tools and services available. Many struggle. In 2018, the government conducted research on Universal Credit online submissions and found that more than half (54%) of all claimants were unable to register their claim online unassisted, 21% of claimants needed help to complete their online application, while 25% were unable to submit their claim online at all.[149] One claimant struggling with the DWP's 'digital by default' system told Amnesty International: "I need to tell them over the internet… all my expenses and all my income for the previous month. It can be frustrating. I'm not saying they try to make it difficult for you, but you go on the internet, and you type in everything and then for whatever reason, they say, 'Oh, we're going to send you a 6-digit code to continue with this', and send it to your phone. I'm on the computer now anyway. What's the point of sending me this thing? And this is a true thing that happened last week. My phone was broken. So, what do you do from there? Because of that… I cannot fill this in. If I don't fill it in, I'm not going to be paid."[150]

Digital exclusion, whether due to internet access or digital skills, is disproportionately experienced by groups who are already marginalised in society, including elderly, low-income, and disabled people and people without English as a first language. 14% of people in the lowest socio-economic group had no internet access at home compared to only 2% in the highest socio-economic group,[151] and 79% of people living in the UK who didn't use the internet were aged 65 or over.[152] Furthermore, many groups working with refugees and people seeking asylum in the UK have emphasised challenges around digital access among their clients – one survey cites digital access as the most pressing need following housing.[153]

145 Ibid.
146 Office for National Statistic. (2019). Exploring the UK's Digital Divide. https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/articles/exploringtheu ksdigitaldivide/2019-03-04
147 See Ofcom. (2022). Digital Exclusion Review. Ofcom. https://www.ofcom.org.uk/__data/assets/pdf_file/0022/234364/digital-exclusion-review2022.pdf
148 https://www.amnesty.org/en/documents/eur45/9478/2025/en/
149 DWP, Universal Credit: Full Service Survey, June 2018, https://assets.publishing.service.gov.uk/media/5b1a4f9eed915d2cc380163f/universal-credit-full-service-claimant-survey.pdf p13.
150 Amnesty International. (2025). 'Too much technology not enough empathy': How the UK's push to digitise social security harms human rights'. Amnesty International. https://www.amnesty.org/en/documents/eur45/9478/2025/en/
151 See Ofcom. (2022). Digital Exclusion Review. Ofcom. https://www.ofcom.org.uk/__data/assets/pdf_file/0022/234364/digital-exclusion-review2022.pdf
152 Serafino, P. (2019). 'Exploring the Digital Divide'. Office for National Statistics. https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/articles/exploringtheuksdigitaldivide/2019-03-04#how-does-digital-exclusion-vary-with-age-
153 Jama, H. (2023, February 7). Reducing digital exclusion for refugees and people seeking asylum in the UK - Refugee Action. Refugee Action. https://www.refugee-action.org.uk/reducing-digital-exclusion-for-refugees-and-people-seeking-asylum-in-the-uk/#:~:text=Marketing%20Marketing-,Reducing%20Digital%20Exclusion%20For%20Refugees%20And%20People%20Seeking%20Asylum%20In,and%20social%20and%20economic%20exclusion.

Consequently, those who suffer from digital exclusion are also often those who are most in need of filing for social security payments online. The UK House of Lords Economic Home Affairs Committee has recognised this challenge noting that "for some claimants, the [digital] approach is a significant barrier to claiming and managing Universal Credit…Those most affected include people with disabilities, mental or physical health problems, learning disabilities, poor literacy skills, or who do not have English as a first language."[154]

Overall, to comply with human rights law, a social security system must be accessible. However, digitally excluded people are facing significant barriers to accessing social welfare when digital benefit systems are the default mechanism for access. This is having a disproportionate impact on already marginalised groups, preventing people from accessing social security without discrimination. As such, upholding the right to social security in an increasingly digital age will require a huge effort on digital inclusion to widen internet access and digital skills. In this regard, the UK's Digital Inclusion Action Plan[155] represents a positive step towards tackling digital exclusion by bringing together government, local authorities, charities, and industry leaders like Google and Openreach to deliver skills and resources at scale.[156] The innovation fund aims to promote and grow local community projects that help people get online, and a pilot device provision to tackle affordability and literacy barriers.

Additionally, the government has recently announced that a large-scale digital inclusion drive will accompany the digital ID scheme. This is a positive step towards mitigating the potential implications of 'digital by default' welfare systems on the right to access social security for digitally excluded people.

5.2.2 Automated risk-scoring for fraud detection

Another development of the UK's "digital welfare state" that is impacting people's right to access social security is the development of automated risk-scoring for fraud detection.

The DWP, and other social welfare authorities across Europe,[157] are increasingly introducing AI and algorithmic systems into their operations, such as for assessing claimants' welfare eligibility, or for assessing their risk of welfare fraud to flag them for human investigation. In the UK these tools include the Universal Credit Advances model for predicting whether an advance claim may be fraudulent or not, the Housing Benefit Accuracy Award Initiative (HBAAI) for predicting fraud in relation to housing benefit claims, and various pilot models for predicting fraud related to self-employed earnings, living together, housing and capital.[158] Given the underpinning logic of these systems is broadly the same, their potential impacts on the right to access social security without discrimination are similar, and as such will be discussed together.

If somebody is flagged by one of these algorithmic fraud risk-scoring systems, their name will be sent to local councils who will then investigate their case. Until January 2024, the benefit claims would then be suspended until the investigation was completed, but the decision to suspend benefits claims until the investigation has been completed has since been dropped because of concerns that it could unfairly result in payment delays for legitimate claimants in

154 Serafino, P. (2019). 'Exploring the Digital Divide'. Office for National Statistics. https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/articles/exploringtheuksdigitaldivide/2019-03-04#how-does-digital-exclusion-vary-with-age-

155 Digital Inclusion Action Plan: First Steps. (2025, February 26). GOV.UK. https://www.gov.uk/government/publications/digital-inclusion-action-plan-first-steps/digital-inclusion-action-plan-first-steps

156 Digital Poverty Alliance (2025, April 22). The Digital Inclusion Action Plan: A bold step towards a more Connected UK. https://digitalpovertyalliance.org/news-updates/digital-inclusion-action-plan/

157 For examples in the Netherlands, Serbia, Sweden, France, Denmark, and Spain, see this blog post from the European AI & Society Fund: https://europeanaifund.org/newspublications/how-ai-driven-welfare-systems-are-deepening-inequality-and-poverty-across-europe/

158 Vincent, R. (2025). 'Suspicion by design'. Big Brother Watch, https://bigbrotherwatch.org.uk/wp-content/uploads/2025/07/Suspicion-By-Design-2.pdf

instances where the system may contain inaccuracies.[159] And indeed, accuracy is a concern. For example, the UK's HBAAI profiles each of the nearly one million people receiving housing benefits for their likelihood of committing fraud, with the 400,000 cases perceived by the model to be highest risk sent to councils for review. Over two thirds of those cases were subsequently found to be wrongful flags.[160,161] Meanwhile, being flagged for fraud and taken to investigation is highly stressful for claimants. As part of the investigation of a flag, claimants are required to put together months of documents (such as bank statements, rent books, and payslips) or face suspension of their benefits.[162] The Greater Manchester Coalition of Disabled People (GMCDP) described their anxieties around being investigated as the "fear of the brown envelope", with Rick Burgess of the GMCDP saying: "Disabled people need support – not being ground down by a brutal system that assumes we are fraudulent until proven innocent".[163]

These systems operate in a similar way to the automated decision-making systems identified in the chapter on discrimination (Ch.1); they compare the personal data of individual claimants against swathes of historical DWP data, to identify patterns, and generate a risk score for the claimant. As such, algorithmic fraud detection systems raise the same concerns about algorithmic bias discussed previously; training datasets often reflect existing social biases, leading to problems where bias built into a dataset is reproduced in an algorithmic system's predictions and decisions. This "bias in, bias out" problem has implications for people's right to access social security without discrimination.

The input data used in these systems raises concerns for discrimination not only because of embedded biases, but also because many of the data points are based on protected characteristics to produce a percentage risk score for fraud.[164] The aforementioned HBAAI system, for example, draws on data points such as gender which is a protected characteristic. However, even if protected characteristics are excluded from data sets, other data points, while not protected characteristics themselves, might still be used individually or in combination to infer protected characteristics.[165] For example, the Universal Credit Advances tool references nationality, which could be used to make assumptions about ethnicity, and self-reported illness could be an indicator of disability.

As such, these systems carry high risks for discriminating against protected characteristics either directly, or indirectly through proxy data points.[166] DWP has acknowledged the risks of discrimination arising from their use of AI and algorithmic tools for risk-scoring. In their Fairness Analysis report of the Universal Credit Advances Model, they said there was "statistically significant referral disparity and outcome disparity for the protected characteristics analysed" and acknowledged that "machine learning models designed to assess fraud will inherently have a degree of disparity".[167] Specifically, they found higher disparity rates for age, disability, marriage/civil partnership, as well as nationality. The UK House of Commons Committee of

159 Seddon, P. (2024). 'Universal Credit claims no longer paused while AI fraud checks carried out' BBC. https://www.bbc.co.uk/news/uk-politics-68030762

160 Booth, R. (2024). 'DWP algorithm wrongly flags 200,000 people for possible fraud and error'. The Guardian. https://www.theguardian.com/society/article/2024/jun/23/dwp-algorithm-wrongly-flags-200000-people-possible-fraud-error

161 Vincent, R. (2025). 'Suspicion by design'. Big Brother Watch, https://bigbrotherwatch.org.uk/wp-content/uploads/2025/07/Suspicion-By-Design-2.pdf

162 Vincent, R. (2025). 'Suspicion by design'. Big Brother Watch, https://bigbrotherwatch.org.uk/wp-content/uploads/2025/07/Suspicion-By-Design-2.pdf

163 Disability Rights UK. (No Date). 'DWP urged to reveal algorithm that 'targets' Disabled people for benefit fraud'. Disability Rights UK. https://www.disabilityrightsuk.org/news/2021/november/dwp-urged-reveal-algorithm-%E2%80%98targets%E2%80%99-disabled-people-benefit-fraud?srsltid=AfmBOorTmXPeUmfskZf2pApQ_FzxUxyC79AOXWzdSBrsNbmAYR_RVDTm

164 Protected characteristics are nine specific attributes legally protected from discrimination, harassment, and victimisation under the UK Equality Act 2010. These include age, disability, gender reassignment, marriage/civil partnership, pregnancy/maternity, race, religion/belief, sex, and sexual orientation.

165 https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/lawful-basis/special-category-data/what-is-special-category-data/

166 Vincent, R. (2025). 'Suspicion by design'. Big Brother Watch, https://bigbrotherwatch.org.uk/wp-content/uploads/2025/07/Suspicion-By-Design-2.pdf

167 UK Government. (2024).'Advances Model Fairness Analysis Summary Report'. Department for Work and Pensions. https://www.whatdotheyknow.com/request/ai_strategy_information/response/2748592/attach/6/Advances%20Fairness%20Analysis%20Feb ruary%2024%20redacted%201.pdf?cookie_passthrough=1

Public Accounts has also raised concerns about "the potential negative impact on protected groups and vulnerable customers of DWP's use of machine learning to identify potential fraud".[168] In 2021, the legal firm Foxglove worked with the GMCDP to challenge the DWP over its use of one of their algorithms, which they believe targeted disabled people in a discriminatory way.[169] Rights groups across Europe have raised concerns about similar tools in welfare settings discriminating against migrants, minority ethnic communities, women, and people with disabilities.[170]

Within digital rights circles, there are debates about whether risk-scoring tools such as those used by the DWP can be made 'unbiased' – such as through increased monitoring and evaluation, and accordingly adjusting data model weightings to compensate for any identified bias. For example, officials in Amsterdam made claims in 2023 that they were building a "fair" algorithm called Smart Check to detect welfare fraud that complied with a framework of technical and ethical guidelines meant to ensure fairness.[171] But when the model was evaluated, despite adjustments to compensate for identified bias, it was found to disproportionately flag welfare applicants with children, women and Dutch nationals. While they had managed to remove previous model bias against people with a migration background, other forms of discrimination persisted. Amsterdam city officials subsequently scrapped the model altogether. Despite the effort the city put in to recalibrate the model and avoid bias – which cost over €500,000 – it still led to discrimination.[172] This raises the question as to whether algorithmic and AI tools for calculating benefit claimants' risk of fraud, including those used by the DWP, can ever be made fair. As such, questions remain as to whether these tools can ever be compatible with our right to social security, which must be realised through equal access. For this reason, many digital rights advocates are calling for prohibitions on AI and algorithmic tools within welfare settings because of the unacceptable risks of discrimination that they pose.

CHAPTER 5 SUMMARY POINTS:

- The right to social security obliges states to ensure that citizens can access benefits when they are unable to work (whether due to old age, unemployment, sickness, or caring for dependencies) and that everybody can access health services.
- In the digital age, 'digital by default' benefit systems and automated decision-making systems for accessing welfare raise concerns for people's ability to access social security without discrimination.
- Digital exclusion means that people may not be able to access 'digital by default' benefit systems, thereby infringing on their right to social security.

168 House of Commons Committee of Public Accounts, DWP Customer Service and Accounts 2023-24 Sixth Report of Session 2024–25 HC 354, 31 January 2025
169 Foxglove. (2021). ' secret algorithm targets disabled people unfairly for benefit probes – cutting off life-saving cash and trapping them in call centre hell'. Foxglove. https://www.foxglove.org.uk/2021/12/01/secret-dwp-algorithm/
170 Banerji, L. and Cabo, D. (2025). 'How AI-driven welfare systems are deepening inequality and poverty across Europe'. European AI and Society Fund. https://europeanaifund.org/newspublications/how-ai-driven-welfare-systems-are-deepening-inequality-and-poverty-across-europe/
171 Guo, E. Geiger, G., Braum, JC. (2025). 'Inside Amsterdam's high-stakes experiment to create fair welfare AI'. MIT Technology Review https://www.technologyreview.com/2025/06/11/1118233/amsterdam-fair-welfare-ai-discriminatory-algorithms-failure/
172 Ibid.

- Automated decision-making or decision-support systems in welfare settings may encode discrimination because they unavoidably rely on protected characteristics to determine outcomes, which could infringe on their right to access social security without discrimination.
- Research shows that attempts to 'debias' automated decision-making systems in welfare settings, are often unsuccessful with other forms of algorithmic discrimination persisting.
- Protecting the right to social security in the digital age will require 'digital benefits systems' to be accompanied by equal non-digital alternatives and strong government digital inclusion initiatives. It may also require prohibitions on automated decision-making in welfare settings.

CHAPTER 6
RIGHT TO WORK IN THE DIGITAL AGE

The right to work is enshrined in Article 23 of the UDHR:[173]

1. *Everyone has the right to work, to free choice of employment, to just and favourable conditions of work and to protection against unemployment.*
2. *Everyone, without any discrimination, has the right to equal pay for equal work.*
3. *Everyone who works has the right to just and favourable remuneration ensuring for himself and his family an existence worthy of human dignity, and supplemented, if necessary, by other means of social protection.*
4. *Everyone has the right to form and to join trade unions for the protection of his interests.*

Article 7 (a-d) of the ICESCR similarly recognizes *"the right of everyone to the enjoyment of just and favourable conditions of work which ensure, in particular, … fair wages and equal remuneration for work of equal values,... a decent living for themselves and their families,… safe and healthy working conditions,... equal opportunity for promotion,... rest, leisure and reasonable limitation of working hours."*[174]

Meanwhile Article 8 of the ICESCR specifies *"the right of everyone to form trade unions and join the trade union of his choice…for the promotion and protection of his economic and social interests."*[175]

173 United Nations General Assembly. (1948). Universal declaration of human rights (217 [III] A). https://www.un.org/en/about-us/universal-declaration-of-human-rights
174 United Nations General Assembly. (1966). International Covenant on Economic, Social and Cultural Rights. United Nations, Treaty Series, vol. 993, p. 3. https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-economic-social-and-cultural-rights
175 United Nations General Assembly. (1966). International Covenant on Economic, Social and Cultural Rights. United Nations, Treaty Series, vol. 993, p. 3. https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-economic-social-and-cultural-rights

6.1 ABOUT THE RIGHT TO WORK

The right to work is a fundamental human right that is intended to protect everybody's right to continue their life with a desired job and an income that allows for the protection of their human dignity. The right to work is included in international and regional human rights frameworks including the UDHR, the ICESCR. In the UK, the ECHR and the HRA do not explicitly protect the right to work, though some articles of the ECHR can be used to protect aspects of the right to work. For example Article 11 on freedom of association and assembly protects the right to join and form trade unions.[176]

Aspects of the right to work may also be protected by domestic employment law which also offers protections to workers that may go beyond human rights law. In the UK, employment law covers aspects such as discrimination, bullying and harassment, dismissals, grievances, contracts, pay, wages, leave, and redundancy.[177] Most recently, the Employment Act 2025 introduces a raft of new workers' rights in the UK such as improved pregnancy, maternity and paternity rights and day one sick leave; curbs on exploitative zero-hours contracts and unscrupulous fire and rehire practices; and greater protection from unfair dismissal.[178]

6.2 IMPACTS OF DIGITAL TECHNOLOGY ON THE RIGHT TO WORK

The rise of new technology is rapidly transforming work and working conditions across the digital value chain. These changes are taking place at both individual and structural levels of the labour market with significant impacts on the right to work, as enshrined within the UDHR and the ICESCR.

At an individual level, AI, algorithms and data-driven technologies are increasingly being used to mediate relationships between employees and their employers – such as for monitoring and surveilling employees, or for automating key managerial decisions such as hiring, firing, and pay allocation decisions.

At a structural level, the emergence of AI is profoundly restructuring the labour market and the types of employment opportunities available to workers. The emergence of the platform economy which, encompassing a range of economic and social activities facilitated by digital platforms such as Uber, Airbnb, Deliveroo, and JustEat, now employs a large proportion of the UK workforce.

This section brings together examples illustrating how these individual and structural changes are having impacts on two aspects of the right to work in particular: the right to fair wages and equal remuneration, and the right to form and join trade unions. This section also provides an overview of how the increasing adoption of AI, algorithms and digital technologies in the workplace hold particular challenges for workers' right to privacy. While the right to privacy in the digital age is covered in greater depth in Chapter 2, we consider it again here because of the increasing prevalence of workplace surveillance.

The end of the chapter offers an overview of the Trades Union Congress' AI and Employment Rights Bill which, as a "ready to go" draft legislation, showcases the rights and obligations necessary to protect workers' rights in light of the increasing adoption of AI and algorithms in the workplace.

176 Gutterman, A. (2024). Rights at Work: Labor Practices and Human Rights. https://pure.qub.ac.uk/en/publications/the-right-to-work-in-the-echr/

177 Key employment legislation in the UK includes: Employment Rights Act 1996; National Minimum Wage Act 1998; Employment Relations Act 1999; The Equality Act 2010; and Health and Safety at Work etc. Act 1974.

178 Matheou, D. (2025, December 16). Parliament passes the Employment Rights Bill. https://www.unison.org.uk/news/2025/12/parliament-passes-the-employment-rights-bill/

6.2.1 Fair wages and equal remuneration

The emergence of AI and digital technologies for calculating wages poses challenges to the right to fair wages and equal remuneration. Of particular concern is the increasingly prevalent use of dynamic pricing algorithms to automatically set variable pay. Dynamic pricing algorithms use AI and machine learning methods to analyse data – such as inventory levels, time of day, and browsing history – to instantly change prices. They are increasingly used by platform employers such as Uber and Deliveroo.[179]

They threaten the right to fair wages because they do not allow workers to anticipate their earnings before they begin their work – instead they are subject to unpredictable (and often opaque) decision-making procedures of an algorithm determining what pay they will receive at a given time. The ability to anticipate earnings is considered to be a core pillar for upholding fair wages to prevent the precarity, instability and stress arising from unpredictable earnings. The UK's Employment Rights Act 1996 requires employers to provide workers with a written statement of employment outlining how much and how often the employee will be paid, and thereby ensuring that the anticipation of earnings is considered an important aspect of fair wages.[180]

Nonetheless, there are growing examples of platform employers such as Uber, Deliveroo, or Lyft using dynamic pricing systems to calculate workers' wages.[181] For example, a groundbreaking algorithmic audit by Worker Info Exchange and the University of Oxford of 1.5 million trips from 258 Uber drivers found that drivers lost 8% in gross annual earnings following the introduction of dynamic pricing algorithms, and that 82% are now earning less per hour than they did before dynamic pay and pricing was introduced. A small minority are earning more, but these gains are not shared evenly with benefits going to newer drivers and part time workers.[182] Indeed, in November 2025 Worker Info Exchange issued Uber with a legal Letter Before Action, demanding the company halt the use of its AI-driven dynamic pay systems.[183] This collection action represents the first in Europe to directly challenge dynamically set, personalised pay, determined through algorithmic decision-making.

6.2.2 Right to form and join trade unions

The right to form, join and participate in trade unions is another pillar of the right to work. It is articulated in both the UDHR and the ICESCR. In the UK, the right to form and join trade unions is also protected in employment law through the Trade Union and Labour Relations (Consolidation) Act 1992 which sets the legal framework for the recognition and functioning of trade unions in the UK, including collective bargaining processes and industrial actions.

However, in the digital age a growing body of workers largely are not able to exercise this right because the protections under the Trade Union and Labour Relations Act do not apply to them. UK employment status creates three distinct categories of workers, who are afforded different rights – employee, worker, and self-employed – and the Trade Union and Labour Relations (Consolidation) Act 1992 only applies to those with 'employee' status.

179 Van Doorn, N. From a Wage to a Wager: Dynamic Pricing in the Gig Economy. Autonomy Institute. https://autonomy.work/wp-content/uploads/2020/09/VanDoorn.pdf

180 Employment Rights Act 1996. https://www.legislation.gov.uk/ukpga/1996/18/contents

181 Gilbert, A. and Marriott, J. (2025). Not just a matter of Oasis tickets: dynamic pricing and algorithms in the workplace. Trust for London. https://trustforlondon.org.uk/news/not-just-a-matter-of-oasis-tickets-dynamic-pricing-algorithms-in-the-workplace/ ; Van Doorn, N. From a Wage to a Wager: Dynamic Pricing in the Gig Economy. Autonomy Institute. https://autonomy.work/wp-content/uploads/2020/09/VanDoorn.pdf

182 Binns, R. et al. (2025). Not Even Nice Work If You Can Get It; A Longitudinal Study of Uber's Algorithmic Pay and Pricing. In Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency (FAccT '25). Association for Computing Machinery, New York, NY, USA, 1484–1497. https://doi.org/10.1145/3715275.3732099

183 Worker Info Exhange. (2025, November 20). Drivers in UK and Europe set to sue Uber for unfair pay set by algorithm. Worker Info Exchange. https://www.workerinfoexchange.org/post/drivers-in-uk-and-europe-set-to-sue-uber-for-unfair-pay-set-by-algorithm

In the digital age, the emergence of the platform economy has brought structural changes in the UK labour market, whereby a growing proportion of the UK workforce undertake temporary, freelance or contract work for online platforms. This means that a growing number of workers do not have their right to form and join trade unions protected in UK employment law. Precise figures on the number of platform workers in the UK are hard to come by, however TUC research in 2021 found 14.7% of working people, estimated at 4.4 million people, were working for gig economy platforms at least once a week in England and Wales. This was based on a survey of 2,201 participants. This has since been extrapolated by StandOut CV, who estimate the size of the gig economy in 2023 as 7.25 million, rising to 14.86 million by 2026.[184]

In 2021, platform workers secured a landmark legal victory when the Supreme Court ruled in the judgement of Uber BV v Aslam that Uber drivers were "workers", and were therefore entitled to holiday leave and the national minimum wage. This was a significant ruling for platform workers' rights with respect to national minimum wage, minimum level of paid holiday, and protection against unlawful discrimination. However their right to trade union representation is still not protected as it is for 'employees' under UK employment law. This means that at least 15% of working people in the UK are not entitled to obtain formal union recognition, nor can they enter into collective bargaining for the platform work they undertake. Many digital rights organisations such as the Institute for the Future of Work (IFOW)[185] and Worker Info Exchange[186] have been calling for a single worker status to ensure that all workers in the UK, including gig economy workers, are able through UK employment law to exercise their right to form, join, and participate in trade unions.

6.2.4 Right to privacy in the workplace

The development and deployment of new technologies in the workplace also carries noteworthy implications for workers' rights to a private life. While the right to privacy is covered in greater detail in Chapter 2, we dedicate additional attention in this section in recognition of the increasing proliferation of workplace surveillance, and the unique implications of privacy for workers.

Workplaces are increasingly adopting digital technologies to mediate relationships between employers and employees for management and monitoring purposes – such as for communications and location monitoring, facial recognition, sensors detecting desk use and worker movements, and technologies for the monitoring of computer activity such as keyboard monitoring and mouse tracking.[187] Businesses including BP, Bank of America, IBM, Target, and Time Warner are known to offer the consumer tech fitness tracker Fitbit to employees as part of corporate wellness programs, with employers able to access dashboards for monitoring employee performance in terms of sleep, activity, and colleague-community fitness challenges.[188] Firms including Toyota, Unilever, P&G and ClearChannel are known to have purchased eye-tracking software, while companies including CenTrak and SwipeSense provide hospitals with systems for assessing staff time management by monitoring how long nurses spend with patients, or aiding hygiene management by assessing how close nurses are to soap dispensers.[189]

184 Cockett, J., Willmott, B., & Chartered Institute of Personnel and Development. (2023). The gig economy: What does it really look like? In Office for National Statistics, Chartered Institute of Personnel and Development [Policy report]. Chartered Institute of Personnel and Development. https://www.cipd.org/globalassets/media/knowledge/knowledge-hub/reports/2023-pdfs/2023-cipd-gig-economy-report-8453.pdf

185 Institute for the Future of Work. Written submission (ULM0082). BEIS Select Committee Inquiry into the UK's Labour Market. https://committees.parliament.uk/writtenevidence/109909/pdf/

186 Worker Info Exchange (2025, June 24). New research exposes deepening exploitation of Uber drivers by algorithmic pay. Worker Info Exchange. https://www.workerinfoexchange.org/post/new-research-exposes-deepening-exploitation-of-uber-drivers-by-algorithmic-pay

187 Atkinson, J. and Evans, J. (2025). Negotiating the future of work: Legislating to protect workers from surveillance. IPPR. https://www.ippr.org/articles/negotiating-the-future-of-work-surveillance

188 Moore et al. (2024). Data on our minds: affective computing at work. Institute for the Future of Work. https://www.ifow.org/publications/data-on-our-minds-affective-computing-at-work

189 Ibid.

Worker surveillance and data tracking enabled by modern digital technology presents a notable threat to privacy due to the extensive personal data collection involved, which may take place disproportionately, unnecessarily, and without an adequate legal basis. Moreover, expert interviewees told us that some workers may not feel they have a choice to decline if they are faced with a choice of handing over their personal data or accessing work. Evidence also suggests that privacy infringements at work may disproportionately affect workers who are already marginalised – for example, a recent report by IPPR found that Black workers in the UK may be at a significantly greater risk of being subjected to workplace surveillance.[190]

TRADES UNION CONGRESS ARTIFICIAL INTELLIGENCE (REGULATION AND EMPLOYMENT RIGHTS) BILL

To advance the regulation of AI in the workplace, the Trades Union Congress has proposed a "ready to go" AI Bill that is a draft law based on a 4 year research project. The TUC AI Bill showcases the rights and obligations necessary to protect workers' rights in light of the increasing adoption of AI and algorithms in the workplace. The Bill includes provisions that, if introduced and adequately enforced, would go significant lengths to alleviating the threats outlined within this chapter on the right to work in the digital age, and especially the impacts of algorithmic decision-making on worker autonomy, algorithmic bias and discrimination, and worker participation in tech adoption in the workplace.

In particular, the TUC AI Bill includes the following provisions that would:[191]

- Prohibit the use of emotion recognition technology that could be detrimental to workers, which would mitigate threats to the right to privacy arising from emotion recognition technology in the workplace.
- Strengthen protections against discriminatory algorithms and shift the burden of proof to employers.
- Strengthen protections against discriminatory algorithms and shift the burden of proof to employers.
- Create a legal duty on employers to consult trade unions before using "high risk" AI in the workplace, which would alleviate the issues of lacking worker consultation or negotiation.
- Give workers the right to a personalised explanation of high-risk decisions made using AI, which would alleviate the issues of lacking transparency which creates challenges for initiating redress processes.
- Establish a right to human review of decisions made by AI systems, which would establish more robust pathways to an effective remedy following tech-induced infringements on workers' rights.

190 Atkinson, J. and Evans, J. (2025). Black employees are at highest risk of being targeted by worker surveillance, report finds. IPPR. https://www.ippr.org/media-office/black-employees-are-at-highest-risk-of-being-targeted-by-worker-surveillance-report-finds
191 Trades Union Congress. (2024). The AI Bill Project. https://www.tuc.org.uk/research-analysis/reports/ai-bill-project

CHAPTER 6 SUMMARY POINTS:

- The right to work covers just working conditions, equal pay and fair remuneration, protection against unemployment, and the right to form and join trade unions.
- In the digital age, AI and data-driven technologies such as dynamic pricing algorithms, firing and hiring algorithms, and monitoring and surveillance technologies raise concerns for the right to work.
- Dynamic pricing algorithms can undermine the right to fair wages and equal remuneration when workers are subject to unpredictable and opaque decision-making procedures to determine pay.
- Platform workers, who make up roughly 15% of UK working people, are excluded from key workers rights that are afforded to 'employees' under the UK's worker status distinctions – including the right to engage in collective bargaining.
- The deployment of AI, algorithms and data-driven technologies in the workplace creates concerns for the rights to privacy and data protection because of unnecessary and disproportionate data processing.
- Protecting the right to work in the digital age will require the creation of a single worker status to ensure all workers, including platform workers, can exercise employment rights. It will also require prohibitions on certain workplace uses of technology, pre-deployment consultations and evaluations, and adequate routes to redress.

CHAPTER 7
RIGHTS TO FREEDOM OF MOVEMENT AND ASYLUM IN THE DIGITAL AGE

The right to freedom of movement is articulated in Article 13 of the UDHR: *"Everyone has the right to freedom of movement and residence within the borders of each state. Everyone has the right to leave any country, including his own, and to return to his country."*[192]

The right to asylum is subsequently articulated in UDHR Article 14: *"Everyone has the right to seek and to enjoy in other countries asylum from persecution. This right may not be invoked in the case of prosecutions genuinely arising from non-political crimes or from acts contrary to the purposes and principles of the United Nations."*[193]

The right to freedom of movement is also reinforced by Article 12 of the ICCPR stating that *"[e]veryone lawfully within the territory of a State shall, within that territory, have the right to liberty of movement"* and *"[e]veryone shall be free to leave any country, including his own."*[194] And the right to asylum has been elaborated on through the 1951 UN Refugee Convention[195] and the 1967 Protocol relating to the status of refugees[196,197] which outline minimum standards of legal protection, rights and assistance a refugee is entitled to receive.

192 United Nations General Assembly. (1948). Universal declaration of human rights (217 [III] A). https://www.un.org/en/about-us/universal-declaration-of-human-rights
193 United Nations General Assembly. (1948). Universal declaration of human rights (217 [III] A). https://www.un.org/en/about-us/universal-declaration-of-human-rights
194 United Nations General Assembly. (1966, December 16). International Covenant on Civil and Political Rights. Treaty Series, 999, 171. https://www.refworld.org/legal/agreements/unga/1966/en/17703
195 Convention Relating to the Status of Refugees, July 28, 1951, 189 U.N.T.S. 137 (entered into force April 22, 1954) https://www.unhcr.org/uk/about-unhcr/overview/1951-refugee-convention
196 United Nations. (1967). Protocol Relating to the Status of Refugees, 606 U.N.T.S. 267 https://treaties.un.org/pages/ViewDetails.aspx?src=IND&mtdsg_no=V-5&chapter=5
197 Asylum seekers and Refugees are two distinct statuses. Asylum seekers are those who have fled their country, seeking protection from persecution but have not yet been recognised as a refugee. Refugees are those who have fled their country, seeking protection from persecution and have been legally recognised as a refugee, making them eligible for international assistance and protection.

7.1 ABOUT THE RIGHTS TO FREEDOM OF MOVEMENT AND ASYLUM

We consider the right to freedom of movement and the right to asylum together because of their combined relevance for migrants and people on the move, who are significantly impacted by the increasing adoption of AI and data-driven technologies in migration contexts. Together the rights require states to ensure individuals can move freely and safely within and across borders, and to seek protection and safety from persecution in their home country.

These rights were initially highlighted on the international stage as governments grappled with the mass displacement of people following the two World Wars. As millions of people were forced to flee across Europe and America, the international community assembled a joint set of guidelines, conventions and laws to protect their basic human rights around movement and asylum. Cementing the right to cross-border travel in international human rights frameworks was a crucial way to pursue such improved international relationships and discourage future conflict.

It is important to note that the rights to freedom of movement and asylum are conditional and ultimately fall subject to the developing laws of respective states. Governments retain the right to restrict and regulate the rights to freedom of movement and asylum in the interest of public security. For example, states across the world restricted travel during the Covid-19 pandemic to limit the spread of the virus.[198]

However, in recent decades there has been a marked change from the more rights-affirming approach governments had towards movement and asylum in the post-war context. The UK has followed this trend. In the face of increasing migration and asylum cases, government policy and public sentiment has grown more hostile.[199] The goal of the current UK government is clear - "net migration must come down."[200] Broader policies proposed by Home Secretary Shabana Mahmood in November 2025 to 'tackle illegal migration' include making refugee status temporary, overhauling human rights law to prevent multiple appeals to asylum cases, and ending housing and financial support for asylum seekers.[201]

The focus on greater restriction and regulation of movement and asylum is also being pursued by the Home Office. Greater digitalisation and deployment of digital technologies is being pursued as part of a move towards a 'fully end-to-end digital' immigration system.[202] The asylum plans announced by Mahmood in November 2025, such as AI-led age checking,[203] coupled with significant expenditure on technological consultants (despite pledges of cost-cutting) confirm that more digital technologies for restricting and regulating movement and asylum are yet to come.[204]

These changes have implications for the rights to movement and asylum, as raised among concerns from digital rights advocates presented in the following section.

198 BBC News. (2020, April 6). Coronavirus: The world in lockdown in maps and charts. https://www.bbc.co.uk/news/world-52103747
199 National Centre for Social Research (2024, June 12). British Social Attitudes. Immigration. https://natcen.ac.uk/publications/british-social-attitudes-41-immigration#:~:text=This%20sharp%20increase%20in%20migration,Central%20African%20state%20of%20Rwanda.
200 HM Government, Prime Minister, & Home Secretary. (2025). Restoring Control over the Immigration System. https://assets.publishing.service.gov.uk/media/6821aec3f16c0654b19060ac/restoring-control-over-the-immigration-system-white-paper.pdf
201 Francis, S. (2025, November 17). Key takeaways: What are the proposed asylum system reforms? BBC News. https://www.bbc.co.uk/news/articles/c3eplk4047do
202 New Plan for Immigration: legal migration and border control (accessible). (2022, November 25). GOV.UK. https://www.gov.uk/government/publications/new-plan-for-immigration-legal-migration-and-border-control-strategy/new-plan-for-immigration-legal-migration-and-border-control-accessible
203 Restoring Order and Control: A statement on the government's asylum and returns policy (accessible). (2025, November 21). GOV.UK. https://www.gov.uk/government/publications/asylum-and-returns-policy-statement/restoring-order-and-control-a-statement-on-the-governments-asylum-and-returns-policy
204 Kundaliya, D. (2025). Home Office increases tech consultant spending to £350m despite cost-cutting pledges. https://www.computing.co.uk/news/2025/home-office-increases-tech-consultant-spending-to-350-million

7.2 IMPACTS OF DIGITAL TECHNOLOGY ON FREEDOM OF MOVEMENT AND ASYLUM

Some uses of technology in migration contexts have had a positive impact facilitating movement, making travel quicker, easier and fuss-free. However, as with the human rights to non-discrimination (Chapter 1), privacy (Chapter 2), effective remedy (Chapter 4), social security (Chapter 5), and work (Chapter 6), the most marginalised and vulnerable people in society disproportionately experience a negative influence of new digital tools on accessing their rights to movement and asylum. A number of migrants' and digital rights organisations have criticised the government's deployment of technology within the immigration system for exacerbating racist and religious discrimination. The examples covered in this chapter show how digital identity systems may cement exclusion[205] and how automated decision-making systems may reinforce discrimination.[206]

Drawing on recent research and interviews from migration, asylum and digital rights experts, this section evaluates some of the key implications of the increasing digitalisation of the immigration system on the right to freedom of movement and right to asylum. We analyse several use cases of digital technology by the Home Office — digital identification systems such as e-Visas, predictive risk assessment systems, and facial age estimations — to illustrate key digital rights concerns.

As the government continues its journey to implement an 'end to end' digital immigration system, the influence on human rights to freedom of movement and asylum, particularly for those most marginalised, must be attended to.

7.2.1 E-Visa digital identification systems

In 2022, the Home Office introduced their 'New Plan for Immigration: legal migration and border control'[207] which set out a vision for a fully digitalised immigration system, characterised by identification systems that are 'digital by default'. In late 2023, e-visas[208] were introduced as part of this new plan, replacing physical identity documents such as the Biometric Residence Permits (BRP), Biometric Residence Cards (BRCs) and legacy documents such as passports. E-Visas are used to demonstrate someone's ability to enter the UK, to prove their right to work, access banking, mortgages or housing agreements.

Considering the centrality of e-Visas for accessing fundamental aspects of daily life, their effective functioning is vital for upholding the rights of the millions of people who are dependent on them. Yet, since their introduction, e-Visas have been widely criticised for undermining migrants' human rights due to their design, functioning and implementation.[209]

Firstly, the roll-out of the mandatory digital-only e-Visa scheme raises concerns pertaining to digital exclusion thereby making people's right to move freely or claim asylum contingent on people's ability to access digital services.[210] The e-Visa sign-up requires access to both stable internet connection and a recent smartphone (iPhone 7 or an Android with contactless

205 Open Rights Group. (2024). E-Visas: Hostile and Broken. https://www.openrightsgroup.org/app/uploads/2025/01/ORG-E-Visa-Report-v2-2UP.pdf
206 Privacy International. (2025). PI alerts regulator about the use of algorithms by the UK Government and their impact on migrants. https://privacyinternational.org/long-read/5639/pi-alerts-regulator-about-use-algorithms-uk-government-and-their-impact-migrants
207 New Plan for Immigration: legal migration and border control (accessible). (2022b, November 25). GOV.UK. https://www.gov.uk/government/publications/new-plan-for-immigration-legal-migration-and-border-control-strategy/new-plan-for-immigration-legal-migration-and-border-control-accessible#planning-to-come-to-the-uk
208 E-vias are defined by the Open Rights Group as 'an online record of a person's immigration status and the conditions of their permission to enter or stay in the UK. It is not an online version of a person's immigration documents but a digital status, generated anew each time it is inspected. People needing to prove their immigration status will need to create a UKVI account to be able to access their e-Visa.' Available at: Open Rights Group. (2024). E-Visas: Hostile and Broken. https://www.openrightsgroup.org/app/uploads/2025/01/ORG-E-Visa-Report-v2-2UP.pdf
209 Ibid
210 Ibid

payment).[211] This presents challenges for migrants and people seeking asylum who often experience barriers to digital connectivity because of limited access to internet and digital devices, language barriers, or poor digital literacy skills.[212] The Home Office suggests that users with limited access to digital devices should borrow the phone of a friend or family member.[213] However, this suggestion presumes that friends or family may be better off than the user themselves, and fails to acknowledge the exploitation such dependency can lead to, especially for young people, women or people with disabilities.[214] In this way, digital exclusion among migrant and asylum-seeking communities present additional challenges for accessing the digital-only e-Visa scheme, and therefore for people to exercise their right to freedom of movement or seek asylum.

Secondly, design flaws and technical glitches of the e-Visa system have also created challenges for people in exercising their right to movement. People have shared that technical glitches have created challenges for them acquiring e-visa's and using them travelling to and from the UK.[215] Irrespective of digital skills, website crashes, login failures, and share code errors have disrupted even the most digitally confident users – in some cases leaving people unable to complete the requisite processes online. One interviewee told researchers that the technical issues with the system created significant challenges for her when travelling back to the UK with her children:

> *"So this year, after making the application for ILR and after the application was approved, they sent us an e-mail saying, "OK, your ILR status has now been approved, you need to link that with your UKVI account" so that wasn't automatically linked so you couldn't log into your account and say, "oh, now this is your new status". You had to link it again, it's such a nightmare to do that because there are different accounts and different things happening, so anyway I managed that and my husband managed that. But when I try to do that for the children. Their account said it expired or that their account didn't exist. So, I tried to do it numerous times, and I'd kept getting different error results, and I wasn't able to link in my children's account to their status."*[216]

Thirdly, reports among the digital rights community have identified that the e-Visa system is prone to errors and glitches because of design flaws. As a live data matching system, e-Visas automatically generate applicants' status by matching names and identities in different databases, and must be regenerated for every inspection. However the system has been found to generate inaccurate information about users, with migrants reporting that as they began to digitalise their statuses and access their eVisas, they noticed that their personal data had become amalgamated with others'.[217] Indeed, an investigation in The Guardian in 2024 found that major technical flaws in the Home Office's immigration database led to more than 76,000 people being recorded with incorrect names, photographs, or immigration statuses which prevented individuals from applying for jobs and housing, accessing services and travel. These technical errors of the e-Visa system have implications for people's ability to exercise their right to freedom of movement because they prevented people from being able to travel.[218]

211 Ibid
212 Jama, H. (2023, February 7). Reducing digital exclusion for refugees and people seeking asylum in the UK - Refugee Action. Refugee Action. https://www.refugee-action.org.uk/reducing-digital-exclusion-for-refugees-and-people-seeking-asylum-in-the-uk/
213 Open Rights Group. (2024). E-Visas: Hostile and Broken. https://www.openrightsgroup.org/app/uploads/2025/01/ORG-E-Visa-Report-v2-2UP.pdf
214 Ibid.
215 Open Rights Group (2025). Exclusion by Design. https://www.openrightsgroup.org/app/uploads/2025/12/Exclusion-by-Design-Report.pdf
216 Ibid.
217 Open Rights Group. (2024). E-Visas: Hostile and Broken. https://www.openrightsgroup.org/app/uploads/2025/01/ORG-E-Visa-Report-v2-2UP.pdf
218 Open Rights Group (2025). Exclusion by Design. https://www.openrightsgroup.org/app/uploads/2025/12/Exclusion-by-Design-Report.pdf

Nonetheless, despite the numerous risks and faults of the e-Visa system, there are no published contingency plans for situations where e-Visas fail. In the terms and conditions of e-Visas, the Home Office has stated that they take no liability for any problems, disruptions or direct or indirect losses when using a UKVI account.[219] The lack of a non-digital alternative or route to appeal leaves people vulnerable to losing access to their right to movement as a result of digital identification systems such as the e-Visa.[220]

7.2.2 Predictive risk assessment technologies

As part of its goals to move towards a 'fully end-to-end digital' immigration system, the Home Office has also been increasingly designing and deploying predictive risk assessment technologies in the visa and asylum process with the purpose of streamlining decision-making, filling information gaps, and supporting the maintenance of a compliant environment.[221] Examples of individual risk assessment technologies that are being used in the UK immigration system include algorithms for determining 'sham marriages', identifying and prioritising cases for immigration enforcement such as detention or removal, and evaluating whether an individual should remain subject to an ankle tag.[222] In logic, these risk assessment systems function similarly to the crime prediction systems discussed in Chapter 1 on equality and non-discrimination, and the fraud detection algorithms covered in Chapter 5 on the right to social security. As such, they raise similar concerns for algorithmic bias, undermining people's rights to freedom of movement and asylum without discrimination.

Digital rights advocates have raised concerns that the 'sham marriage' risk prediction system may discriminate on the basis of ethnicity or nationality.[223] The algorithm, used since at least April 2019, triages applicants into green and red categories according to pre-determined risk factors. Those who are allocated a red rating receive further scrutiny from Home Office officials, and may be investigated further through interviews, or house visits.[224] The algorithm draws on eight risk factors from which three have been publicly disclosed including the age difference between partners, shared travel events, and the observations made by the registrar. The remaining five risk factors are not available to public knowledge. Whilst the 'risk factors' used by the algorithm do not include protected characteristics, documents from the Home Office display some nationalities – including Bulgarian, Greek, Romanian and Albanian – are disproportionately flagged as 'Red' more than other nationalities, which suggests the algorithm may discriminate on the basis of proxies for protected characteristics. In response the Public Law Project (PLP) launched legal action against the Home Office in 2023 based on concerns that the automated decision-making system being used to flag 'sham' marriages was discriminatory with calls to fully disclose the 'risk factors' used.[225] In this way, the potentially discriminatory risk factors of predictive algorithms in migration contexts may have implications for people's ability to exercise their right to freedom of movement and to do so without discrimination.

219 UKVI account: terms and conditions. (2024, February 19). GOV.UK. https://www.gov.uk/government/publications/ukvi-account-terms-and-conditions/ukvi-account-terms-and-conditions#exclusion-of-liability
220 Ibid.
221 Privacy International. (2025). PI alerts regulator about the use of algorithms by the UK Government and their impact on migrants. https://privacyinternational.org/long-read/5639/pi-alerts-regulator-about-use-algorithms-uk-government-and-their-impact-migrants
222 Ibid.
223 Ozkul, D. & Algorithmic Fairness for Asylum Seekers and Refugees (AFAR) Project. (2023). Automating Immigration and Asylum: The uses of new technologies in migration and asylum governance in Europe [Report]. Refugee Studies Centre, University of Oxford. https://www.rsc.ox.ac.uk/files/files-1/automating-immigration-and-asylum_afar_9-1-23.pdf
224 Ozkul, D. & Algorithmic Fairness for Asylum Seekers and Refugees (AFAR) Project. (2023). Automating Immigration and Asylum: The uses of new technologies in migration and asylum governance in Europe [Report]. Refugee Studies Centre, University of Oxford. https://www.rsc.ox.ac.uk/files/files-1/automating-immigration-and-asylum_afar_9-1-23.pdf
225 The case made by the Public Law Project was dismissed in April 2024. See: HM Courts & Tribunals Service. (2024, April 18). Public Law Project v Information Commissioner: [2024] UKUT 71 (AAC). GOV.UK. https://www.gov.uk/administrative-appeals-tribunal-decisions/public-law-project-v-information-commissioner-2024-ukut-71-aac

7.2.3 AI-driven Facial Age Estimation (FAE)

Digital rights advocates also raised concerns about the government's announcement to introduce AI-enabled Facial Age Estimation into UK immigration contexts to verify the age of unaccompanied asylum-seeking children and help determine if adults are falsely claiming to be children. However, inaccuracies of automated age estimations may implicate people's ability to seek asylum, meaning that children are wrongfully flagged as adults, and subsequently housed with adults or detained.[226]

In early 2025 the Minister for Border Security and Asylum stated "...we have concluded that the most cost-effective option to pursue [to ensure that adults are not wrongly identified as children] is likely to be facial age estimation, whereby AI technology trained on millions of images where an individual's age is verifiable is able to produce an age estimate with a known degree of accuracy for an individual whose age is unknown or disputed."[227] Using AI to verify the age of unaccompanied asylum-seeking children has been reaffirmed in the 2025 asylum overhaul led by Shabana Mahmood.

However, the accuracy of AI age checks has been broadly disproved[228] and its use in asylum contexts has been criticised by migrants' rights organisations for additional discrimination and privacy concerns, and leading to examples where children have been wrongfully treated as adults within the asylum system. Indeed, reports indicate that a significant number of children (over 1,300 between Jan 2022 and June 2023) were wrongly classified as adults.[229] Organisations such as Right to Remain point out Facial Age Estimation technology cannot take into account the impact of experiences such as refugee camps, extreme grief, sun exposure and trauma on the appearance of a child, which may lead the technology to estimate that the child is older than they are.[230]

Additionally, The Independent Chief Inspector of Borders and Immigration (ICIBI), who is an independent body that inspects immigration and asylum procedures, highlighted that the Home Office's use of age assessments between July 2024 - February 2025 created a culture of disbelief in a child's stated age, and amplified racial bias whereby Facial Age Estimation Technology has been proven to disproportionately wrongly classify racially minoritised children as adults.[231,232]

The consequences of a wrong age estimation can be dire by placing the child at risk of exploitation in adult facilities, in detention, or even returning them to the country in which they are at risk.[233] The use of AI age assessments in this way thus greatly threatens the right of vulnerable children to safely seek and be granted asylum.

226 https://hansard.parliament.uk/commons/2025-07-22/debates/25072227000021/IndependentChiefInspectorOfBordersAndImmigrationReportAgeAssessmentChecks ; Taylor, D. (2024, January 23). More than 1,000 child refugees at risk after being classified as adults – report. The Guardian. https://www.theguardian.com/uk-news/2024/jan/22/flawed-age-assessments-put-hundreds-of-uk-child-refugees-at-risk-report-finds-home-office
227 https://hansard.parliament.uk/commons/2025-07-22/debates/25072227000021/IndependentChiefInspectorOfBordersAndImmigrationReportAgeAssessmentChecks
228 Asare, J. G., PhD. (2025, August 13). AI Age Checks Are Here—And they're not fair to everyone. Forbes. https://www.forbes.com/sites/janicegassam/2025/08/13/ai-age-checks-are-here-and-theyre-not-fair-to-everyone/
229 Taylor, D. (2024, January 23). More than 1,000 child refugees at risk after being classified as adults – report. The Guardian. https://www.theguardian.com/uk-news/2024/jan/22/flawed-age-assessments-put-hundreds-of-uk-child-refugees-at-risk-report-finds-home-office
230 Artificially intelligent, genuinely harmful: AI and age assessments in the UK asylum system. (2025, August 7). Right to Remain. https://righttoremain.org.uk/artificially-intelligent-genuinely-harmful-ai-and-age-assessments-in-the-uk-asylum-system/
231 Immigration, I. C. I. O. B. A. (2025, July 22). An inspection of the Home Office's use of age assessments (July 2024 – February 2025). GOV.UK. https://www.gov.uk/government/publications/an-inspection-of-the-home-offices-use-of-age-assessments-july-2024-february-2025
232 Asare, J. G., PhD. (2025, August 13). AI Age Checks Are Here—And they're not fair to everyone. Forbes. https://www.forbes.com/sites/janicegassam/2025/08/13/ai-age-checks-are-here-and-theyre-not-fair-to-everyone/
233 Taylor, D. (2024, January 23). More than 1,000 child refugees at risk after being classified as adults – report. The Guardian. https://www.theguardian.com/uk-news/2024/jan/22/flawed-age-assessments-put-hundreds-of-uk-child-refugees-at-risk-report-finds-home-office

CHAPTER 7 SUMMARY POINTS:

- The right to freedom of movement protects everybody's right to move freely within borders and to leave any country. The right to asylum protects people's ability to seek safety in other countries.
- In the digital age, the deployment of technologies in the UK immigration system raises concerns for the rights to freedom of movement and asylum, including: e-visa systems, predictive risk assessment technologies, facial age estimation technologies, and AI-enabled border surveillance.
- The e-visa system may create barriers for digitally excluded people, and design flaws, errors and inaccuracies may prevent people from exercising their right to freedom of movement.
- Predictive risk assessment technologies may embed algorithmic discrimination, undermining people's right to freedom of movement and asylum without discrimination.
- Inaccuracies arising from AI-driven Facial Age Estimation, which is known to have high error rates, may lead children to be wrongfully treated as adults, undermining their right to seek asylum.
- Protecting the rights to asylum and freedom of movement in the digital age will require ensuring everyone has access to equal non-digital systems, robust procedural safeguards, and prohibitions on certain technology use cases that are incompatible with fundamental rights.

PART 2
RECOMMENDATIONS FOR UPHOLDING DIGITAL RIGHTS

This report has illustrated the wide-ranging challenges that AI and other new and emerging data-driven digital technologies can pose to fundamental human rights. As the public continues to experience both the opportunities and challenges of technology, it is crucial that people trust that their interests are at the heart of the government's regulatory approach, and that people feel the state is on their side.

Amidst largely deregulatory approaches to the governance and regulation of technology and AI, and without adequate legislative and regulatory frameworks to safeguard people's rights in the digital age, the UK government risks facing ongoing resistance and public backlash to its plans for technological transformation. The result of this approach is a further erosion of trust that is needed between citizens and state for a well-functioning democratic society more broadly.

At Demos, we believe that digital rights are fundamental for rebuilding the relationship between citizen and state in the digital age. We believe that the government urgently needs a new deal for technology, by placing human rights at the centre of its considerations, and using human rights standards and law to guide new digital policy, regulation, and government deployment of tech solutions.

In this final section we offer recommendations for grounding a rights-based approach to digital policy and tech adoption in the UK, and embedding a new deal for technology within UK policymaking.

1. A UK DECLARATION ON DIGITAL RIGHTS

The UK government should make a firm commitment to upholding citizens' human rights in the digital age, by supporting the development of and signing onto a Declaration on Digital Rights.

A UK Declaration on Digital Rights would be a principle-based document rooted in international human rights frameworks. As a powerful symbol, a declaration would act as a statement of government commitment to protect non-negotiable human rights in the face of technological change. Practically, the declaration would offer high-level guidance and normative foundations for the UK government's approach to tech regulation, and would provide a guide rail for the UK's digital policy development and a checklist for future tech regulation.[234] Indeed, in other countries, similar initiatives such as the EU's Declaration on Digital Rights and Principles (2022) and South Korea's Digital Bill of Rights (2023) have subsequently been followed by legally-binding regulation intended to uphold digital human rights, namely the EU AI Act and South Korea's AI Basic Act.

As such, a UK declaration could set a valuable precedent for guiding future tech policy and regulation. It would serve a 'start point' function, providing a base on which more concrete policy and regulatory work can develop, as well as a valuable 'cultural influence' function; principles provide a common framework for contemplating the potential impacts and risks posed by digital technologies. In turn, those principles influence the kinds of solutions considered, and overtime can help embed cultural norms around tech regulation and human rights. Amidst a growing trend of "digital deregulation", the digital rights declaration could set a direction for tech regulation that puts human interests first.

We provide a draft Declaration for the UK in the Appendix of this report to illustrate the kind of rights-based document that the UK government might commit to. This example Declaration was developed by the Demos internal team, and informed by expert stakeholder consultations. The draft Declaration should therefore not be considered as a final product but a start point for further development. Further development of the Declaration should continue to engage civil society organisations and impacted communities and their advocates organisations in parity with other stakeholders, such as businesses and state representatives.

As a declaration of this sort must ultimately be translated into practice to have concrete impact, we offer the following four recommendations as steps the government should take to deliver on a declarative commitment to digital rights:

234 Seger, E. (2022). In Defence of Principlism in AI Ethics and Governance. Philos. Technol. 35, 45. https://doi.org/10.1007/s13347-022-00538-y

2. BINDING AND ENFORCEABLE HUMAN RIGHTS-BASED TECH REGULATION

We recommend that the UK government introduces robust legal frameworks and safeguards on technology use cases in all sectors to respect, protect, and promote human rights as directed by a public declaration on digital rights.

These regulations should prevent harmful technology systems from being used, or require decision-makers to consider and address the threats to human rights that might arise. This is particularly important for public sector services, and especially criminal legal, migration, and military contexts, where disproportionate, unaccountable, inaccurate or discriminatory technology use cases could significantly impact a person's life.

Binding and enforceable tech regulation should include: ongoing Human Rights Impact Assessments (HRIAs), Data Protection Impact Assessments (DPIAs), Equality and Community Impact Assessments (EIAs and CIAs) throughout the lifecycle of technology development and deployment; mandatory human rights due diligence for businesses; and mandatory information disclosure mechanisms.

As part of the UK government's binding and enforceable human rights-based tech regulation, we recommend that they introduce horizontal AI regulation setting legally-binding obligations for the design, deployment and ongoing evaluation of AI and digital technologies across all sectors by state, non-state, and private actors:

- This regulation should be horizontal, applicable to all sectors while not precluding possibilities of additional regulation for specific sectors.
- The regulation should be adaptive to ongoing technological changes, as regulatory frameworks that are unable to maintain pace with rapid technological innovation hold risks for human rights because they may quickly become outdated, and therefore obsolete.
- The regulation should apply to both public and private sectors. This report has shown that corporate actors hold significant responsibility for many of the risks to human rights arising from AI and digital technologies. As such, the UK's AI regulation must ensure that regulatory authorities can hold corporations to account for any actions that may undermine fundamental human rights, democracy, or the rule of law.
- The regulation should not include blanket exemptions for certain public sector services – and especially those where disproportionate, unaccountable, inaccurate or discriminatory technology use cases could significantly impact a person's life, such as criminal legal, migration, and military contexts.

Finally, the UK government should ensure that supervisory authorities across sectors – including regulators, consumer protection authorities, and judicial review entities – possess sufficient resources and expertise to effectively monitor potential rights infringements arising from technology use cases and decisively intervene to enforce regulatory requirements.

3. REDLINES ON UNACCEPTABLE USE CASES

The UK government should introduce redlines against the development, deployment, import and export of technology applications by state, non-state and private actors that present unacceptable risks to fundamental rights.

During the 80th session of the United Nations General Assembly in September 2025, a broad group of prominent leaders in policy, academia and industry, including ourselves at Demos, launched a call for an international agreement on redlines for AI – ensuring they are operational, with robust enforcement mechanisms – by the end of 2026. The international redlines to prevent unacceptable AI risks and protect against the increasing dangers of AI to human rights and society.[235]

Among UK policymakers, politicians, civil society, and the public, there have been calls for specific prohibitions on the following technologies:

- Remote or retrospective biometric identification (RBI) including Live Facial Recognition (LFR) and emotion recognition technology;[236]
- Crime prediction systems (both individual and geographic);[237]
- Spyware;[238]
- Individual risk prediction algorithms and AI systems in welfare[239] and immigration settings.[240]

In response to UK and international calls, the government should establish redlines accordingly on technology use cases presenting unacceptable risks to human rights.

There is significant emphasis among the digital rights community on the need to introduce these redlines as total bans without exemptions for public authorities such as law enforcement, border, national security, and counterterrorism authorities where uses of AI, algorithms, and data-driven technology can have significant impacts on a person's life.[241]

235 AI Red Lines. (2025). 200+ prominent figures endorse Global Call for AI Red Lines. https://red-lines.ai/#call

236 Big Brother Watch. (2023). 65 parliamentarians call for "immediate stop" to live facial recognition surveillance. https://bigbrotherwatch.org.uk/press-releases/65-parliamentarians-call-for-immediate-stop-to-live-facial-recognition-surveillance/ ; Big Brother Watch. (2023). 180+ tech experts call for global stop to facial recognition surveillance. https://bigbrotherwatch.org.uk/press-releases/180-tech-experts-call-for-global-stop-to-facial-recognition-surveillance/ ; Trades Union Congress. (2024). Artificial Intelligence (Regulation and Employment Rights) Bill. https://www.tuc.org.uk/research-analysis/reports/artificial-intelligence-regulation-and-employment-rights-bill

237 Skelton, S. K. (2025, June 27). MPs propose ban on predictive policing. ComputerWeekly.com. https://www.computerweekly.com/news/366626658/MPs-propose-ban-on-predictive-policing; Big Brother Watch et al. (2025, March 2). Joint letter. Law enforcement use of Automated Decision-making. https://www.statewatch.org/media/4874/uk-law-enforcement-adm-letter-21-3-25.pdf ; Amnesty UK. (2026, January 7). 19,403 people called on the UK to band "crime predicting" technology. https://www.amnesty.org.uk/19000-people-called-uk-ban-crime-predicting-tech

238 United Nations Human Rights Office of the High Commissioner. (2021). Spyware scandal: UN experts call for moratorium on sale of 'life threatening' surveillance tech. https://www.ohchr.org/en/press-releases/2021/08/spyware-scandal-un-experts-call-moratorium-sale-life-threatening

239 Amnesty International. (2025, July 10). UK: Government's unchecked use of tech and AI systems leading to exclusion of people with disabilities and other marginalized groups. https://www.amnesty.org/en/latest/news/2025/07/uk-governments-unchecked-use-of-tech-and-ai-systems-leading-to-exclusion-of-people-with-disabilities-and-other-marginalized-groups/

240 Amnesty International. (2025b, September 12). Advocacy briefing for defending the rights of refugees, asylum seekers, and migrants in the digital age - Amnesty International. https://www.amnesty.org/en/documents/pol30/0290/2025/en/?trk=feed_main-feed-card_feed-article-content

241 Perry, H. et al. (2025). ADVANCING DIGITAL RIGHTS IN 2025 TRENDS, CHALLENGES, AND OPPORTUNITIES IN THE UK, EU AND GLOBAL LANDSCAPE. Demos. https://demos.co.uk/wp-content/uploads/2025/02/Digital-Rights-in-2025.ac_.pdf

4. TRANSPARENCY, ACCOUNTABILITY AND REDRESS

The UK government should ensure that all uses of technology are developed and deployed with robust and meaningful transparency measures to support accountability and redress.

Robust transparency measures are a minimum first requirement for ensuring that individuals and communities are able to initiate redress processes in cases where state, non-state and private actors may have infringed on their rights. Meaningful transparency measures must be evaluated according to their ability to support people to initiate redress processes, apply across different

stages of the digital value chain, and include obligations for alerting people to possible impacts on their rights and ensuring they are able to evidence those harms.

For the UK, significant progress would be made on transparency for public sector organisations' uses of AI, algorithms, and digital technologies by placing the Algorithmic Transparency Recording Standard (ATRS) on a statutory footing, as outlined in the chapter on the right to an effective remedy, to strengthen legal obligations for public sector authorities to disclose what systems they are using.[242] We understand that DSIT is currently looking into improving the system.

Examples of transparency legislation that have been introduced as part of AI governance initiatives in other jurisdictions include: the EU's AI Act, California's SB53, New York's RAISE Act, and the Republic of Korea's AI Basic Act. The EU's AI Act requires providers of high-risk AI systems to provide information to deployers so that decisions of the AI system can be explained to users, and to inform users of certain AI systems, including general-purpose models, that they are interacting with an AI system.[243] California's SB53 and New York's RAISE Act both require large frontier developers to publish a safety and security framework, including policies for evaluating 'critical' or 'catastrophic' AI risks.[244] Meanwhile, the Republic of Korea's AI Basic Act requires AI operators to provide content notices for AI-generated content, and explainability requirements for 'high impact' AI systems on the outcomes, key criteria and principles used in such an outcome, and a summary of the AI's training data.[245]

Finally, transparency obligations should not exempt public sector authorities such as law enforcement, border, national security, or counterterrorism authorities. In settings where potential rights infringements can have serious consequences for somebody's life, it is all the more crucial that people have access to information about AI and tech systems so they can seek accountability and redress.

242 Public Law Project. (2023). Securing meaningful transparency of public sector use of AI. Comparative approaches across five jurisdictions. https://publiclawproject.org.uk/content/uploads/2024/10/Securing-meaningful-transparency-of-public-sector-AI.pdf
243 Key issue 5: Transparency obligations. EU AI Act. https://www.euaiact.com/key-issue/5
244 California Senate Bill no.53. (2025). Artificial intelligence models: large developers. https://legiscan.com/CA/text/SB53/2025 ; New York Senate Bill A6953A. (2025). Responsible AI safety and education act. https://www.nysenate.gov/legislation/bills/2025/S6953/amendment/A
245 Republic of Korea, AI Basic Act (2025). https://www.law.go.kr/LSW/eng/engLsSc.do?menuId=2&query=FRAMEWORK%20ACT%20ON%20THE%20DEVELOPMENT%20OF%20ARTIFICIAL%20INTELLIGENCE%20AND%20THE%20CREATION%20OF%20A%20FOUNDATION%20FOR%20TRUST#liBgcolor0

5. MEANINGFUL PUBLIC PARTICIPATION IN TECHNOLOGY AND AI GOVERNANCE

The UK government should embed meaningful public participation at key points of AI and digital technology policy and governance to ground public trust and facilitate positive technological change.

The UK government should involve citizens meaningfully and equitably at key points in the policy making cycle, to ensure that where there are trade offs to consider, a representative group of citizens are asked to deliberate and reach a shared judgement. Via processes such as citizens assemblies, juries and panels, citizens should be asked to consider the experiences and impacts of technology on the rights of all communities.

Demos' Citizens' White Paper illustrates how putting public deliberation at the heart of policymaking would support the government in navigating complex and divisive policy challenges with greater public support and satisfactory results.[246] Participatory processes such as citizens' assemblies, citizens' juries, co-design workshops, and community conversations to obtain public feedback, manage trade-offs, or ongoing spaces for the public to make recommendations on an issue. Trials of AI-facilitated public deliberation have also enjoyed notable success, allowing for meaningful public deliberation at a larger scale and at significantly lower cost. For example, Pol.is provided a critical role in the aftermath of Taiwan's 2014 sunflower movement for facilitating consensus-based, large-group deliberation to transition from street protests to formal policy-making. Meanwhile, Demos is currently trialling Waves, the largest trial of digital democracy in the UK, to support local councils to engage meaningfully with local citizens to address contentious social issues.[247]

By engaging citizens in tech policy development and governance decisions via participatory methods such as these, the government will be aided in mapping a course through the more complex and divisive debates that have surrounded key tech policy challenges such as AI copyright and digital identification. The policy decisions that result will, in turn, enjoy greater democratic legitimacy, stronger public support for tech transformations, and take a critical step toward rebuilding the badly fractured trust between citizen and state.[248] It is for these reasons that a key recommendation from Ireland's Joint Committee on AI's first interim report (December 2025) is to,"establish a Citizens' Assembly on Artificial Intelligence Digitalisation and Technology to facilitate inclusive public dialogue and democratic input on AI policy and ethics."[249]

A more participatory democracy is at the heart of a new deal to repair the broken relationship between citizens and state, and as Demos CEO Polly Curtis writes, "It is not about ceding power to people, it's about being empowered to represent with renewed trust and legitimacy and to create space for more ambitious policy making. It's about strengthening policy-making and making it more responsive and agile to the scale of the challenges ahead. It's about harnessing the power of state and citizen to move forward together, instead of against one another."[250]

246 Levin et al. (2024). Citizens' White Paper. Demos. https://demos.co.uk/research/citizens-white-paper/

247 Miller, C. (2020, September 27). How Taiwan's 'civic hackers' helped find a new way to run the country. The Guardian. https://www.theguardian.com/world/2020/sep/27/taiwan-civic-hackers-polis-consensus-social-media-platform ; Demos (2025). Waves: Digital Democracy hits the streets of Camden. https://demos.co.uk/blogs/waves-digital-democracy-hits-the-streets-of-camden/

248 Ibid.; Ada Lovelace Institute. (2024). Meaningful public participation and AI. https://www.adalovelaceinstitute.org/blog/meaningful-public-participation-and-ai/ ; Ada Lovelace Institute. (2024). Community-informed governance: reflections for the AI sector. https://www.adalovelaceinstitute.org/blog/community-informed-governance-ai/ ; Ada Lovelace Institute. (2024). Mobilising publics and grassroots organisations to impact AI policy. https://www.adalovelaceinstitute.org/blog/mobilising-publics-impact-ai-policy/

249 Oireachtas, H. O. T. (2025, December 16). Joint Committee on Artificial Intelligence publishes First Interim Report with 85 recommendations. https://www.oireachtas.ie/en/press-centre/press-releases/20251216-joint-committee-on-artificial-intelligence-publishes-first-interim-report-with-85-recommendations/

250 Curtis, P. (2025). Upgrading Democracy: A new deal to repair the broken relationship between citizen and state. Demos. https://demos.co.uk/research/upgrading-democracy-a-new-deal-to-repair-the-broken-relationship-between-citizen-and-state/

APPENDIX
A DRAFT DECLARATION ON DIGITAL RIGHTS

We present this draft **Declaration on Digital Rights** as an example of a human rights-based commitment the UK government might sign onto. The draft Declaration was developed through consultations with more than 40 experts in technology and human rights in the UK and internationally. The incredible breadth of expertise of those we spoke to includes knowledge of: privacy and data protection, platform accountability, redress, state abuses of data, migrants' rights, workers' rights, children's rights, digital inclusion, environmental impacts of AI, and military uses of AI.

From expert interviews, we sought to understand: the main threats to our rights in the digital age, the impacts on technology people's lives, and the legislative and policy recommendations from across the digital rights community. These meetings were followed by workshopping to collaboratively refine sections of the draft declaration.

While our research efforts were extensive, time limitations restricted opportunities for iteration. We therefore present this work as our own and it should not be taken as representative of community consensus.

We offer this draft as an example of a UK government commitment to digital rights. It can be viewed as a starting point for further development which should be undertaken through participatory processes with impacted groups, civil society, and government representatives.

The draft Declaration contains 12 sections. Each section provides a brief interpretation of our fundamental human rights in the digital age, before providing a set of high-level principles for the UK government to uphold our fundamental rights in policy and legislation. Seven of the 12 sections correspond with the chapters in this report, and the accompanying right in international human rights frameworks. The remaining five chapters were not included within the scope of the report because of limitations in time, but were included in the draft declaration in recognition of their urgency in ongoing technology and AI governance conversations. These sections are: corporate accountability & fair competition, children's rights, digital inclusion, the right to a healthy environment, and the right to life, liberty & security of person.

The draft declaration is structured as follows:

1. Equality & non-discrimination
2. Privacy & data protection
3. Freedom of expression, information & assembly
4. Transparency, accountability & redress
5. Corporate accountability & fair competition
6. Adequate standard of living
7. Freedom of movement & asylum
8. Fair & just working conditions
9. Children's rights
10. Digital inclusion
11. Healthy environment
12. Life, liberty and security of person

PREAMBLE

Over the past two decades, digital technology has come to permeate almost every sphere of daily life. Our daily interactions are increasingly mediated by digital technology, algorithms, and AI from accessing essential public services,[251] to communicating with friends and family members,[252] to navigating the workplace.[253]

Some have highlighted the benefits for citizens of this ongoing period of technological change into the era of AI, such as enhanced inclusion, self-expression, productivity, and economic growth.[254] But as with previous periods of technological change, there are a growing number of profound concerns about the economic, social, and environmental impacts brought by the emergence of AI and other data-driven digital technologies.

Governments are increasingly adopting automated surveillance technologies,[255] facial recognition,[256] and biometric identification with concerning implications for fundamental rights to privacy, freedom of movement and asylum, and freedom of expression, information, and assembly. Biased and discriminatory algorithms are being used to make key decisions about people's lives from welfare distribution,[257] hiring and firing decisions,[258] to criminal legal

251 Big Brother Watch (2021). Poverty Panopticon: the hidden algorithms shaping Britain's welfare state. In C. Van Veen & S. Howes, Big Brother Watch. https://bigbrotherwatch.org.uk/wp-content/uploads/2021/07/Poverty-Panopticon.pdf
252 Brunner, L. (2018). Digital communications and the evolving right to privacy. In Cambridge University Press eBooks (pp. 217–242). https://doi.org/10.1017/9781316838952.010
253 Institute for the Future of Work (2025)). Pissarides Reviews. See: https://www.ifow.org/landing-page/the-pissarides-review
254 Knight, S. (2025). Tech that Liberates: A new vision for embedding AI in public service reform. Demos. https://demos.co.uk/research/tech-that-liberates-a-new-vision-for-embedding-ai-in-public-service-reform/
255 Amnesty International UK. (2025). AUTOMATED RACISM: How police data and algorithms code discrimination into policing. https://www.amnesty.org.uk/files/2025-02/Automated%20Racism%20Report%20-%20Amnesty%20International%20UK%20-%202025.pdf
256 Badshah, N. (2025, August 8). Met police to more than double use of live facial recognition. The Guardian. https://www.theguardian.com/technology/2025/jul/31/met-police-to-more-than-double-use-of-live-facial-recognition
257 Big Brother Watch. (2025). Suspicion by Design: What we know about the DWP's algorithmic black box and what it tries to hide. https://bigbrotherwatch.org.uk/wp-content/uploads/2025/07/Suspicion-By-Design-2.pdf
258 Greggwirth. (2025). New study finds AI-enabled anti-Black bias in recruiting - Thomson Reuters Institute. Thomson Reuters Institute. https://www.thomsonreuters.com/en-us/posts/legal/ai-enabled-anti-black-bias/; Worker Info Exchange (2023). Just Eat Report. https://www.workerinfoexchange.org/just-eat-report

proceedings such as policing operations, sentencing decisions, or releases.[259] And large online platforms employ content moderation and recommendation systems that are both limiting access to vital information, and disproportionately amplifying discrimination with tangible offline consequences for people's fundamental rights.[260] And these impacts have a disproportionate negative effect on already marginalized groups, embedding and amplifying offline power structures in a digitally mediated world.[261]

Additionally, private companies have emerged as leading forces in this ongoing technological transformation, accruing dominance over economic markets and political spheres.[262,263]

There is widespread global recognition that policy intervention is needed to ensure citizens can confidently and equitably enjoy the benefits of our ongoing digital transition, knowing that their fundamental rights will not be compromised for the sake of technological progress. The call is ubiquitous throughout civil society, and several governments and international bodies have initiated responses to uphold fundamental rights in the digital age.

In 2023, the European Commission launched the AI Act with the central aim of "developing a strong regulatory framework based on human rights"[264] and the Biden administration released an Executive Order on Artificial Intelligence to "protect Americans' privacy,[265] advance equity and civil rights". Similar initiatives have been launched in Brazil, South Korea, Japan, and South Africa.[266] In 2024, the UN adopted the Global Digital Compact and the Council of Europe adopted the Framework Convention on Artificial Intelligence, becoming the first-ever international legally binding treaty on AI and human rights.[267] The UK is a signatory on both.

Against this backdrop the UK must carve out its own position for securing a positive future for technological development centring on the wellbeing, autonomy, and dignity of people. *Toward this end, signing this Declaration of Digital Rights makes a firm commitment to respecting and protecting people's human rights in the digital age.*

259 Statewatch (2025). New Technology, Old Injustice: Data-driven discrimination and profiling in police and prisons in Europe. https://www.statewatch.org/publications/reports-and-books/new-technology-old-injustice-data-driven-discrimination-and-profiling-in-police-and-prisons-in-europe/

260 Perry, H. and Malik, N. (2025). Researching the riots: An evaluation of the efficacy of Community Notes during the 2024 Southport riots. Demos. https://demos.co.uk/research/researching-the-riots-an-evaluation-of-the-efficacy-of-community-notes-during-the-2024-southport-riots/ ; Amnesty International. (2025). Technical explainer on X's recommender system and the 2024 racist riots. https://www.amnesty.org/en/documents/eur45/0618/2025/en/

261 Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. Proceedings of Machine Learning Research, 81, 77-91. https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf. ; Benjamin, R. (2019). Race after technology: Abolitionist tools for the new Jim code. Polity. ; Noble, S. U. (2018). Algorithms of oppression: How search engines reinforce racism. New York University Press. ; O'Neil, C. (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. Crown.

262 Companies Market Cap (2025). Companies ranked by Market Cap. https://companiesmarketcap.com/gbp/#google_vignette ; Donegan, M. (2025, February 7). It is Elon Musk who is now running the United States. Not Donald Trump. The Guardian. https://www.theguardian.com/commentisfree/2025/feb/06/elon-musk-us-government-power ;Geoghegan, P. (2025, September 24). Inside the Tony Blair Institute. New Statesman. https://www.newstatesman.com/politics/2025/09/inside-the-tony-blair-institute ; Milmo, D. (2025, May 15). Labour's open door to big tech leaves critics crying foul. The Guardian. https://www.theguardian.com/technology/2025/may/14/labours-open-door-to-big-tech-leaves-critics-crying-foul

263 Hao, K. (2025). Empire of AI: Inside the reckless race for total domination. Penguin Books Ltd. ; Srnicek, N. (2026). Silicon empires: The fight for the future of AI. Polity.

264 Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)

265 Executive Order No. 14,110, 88 Fed. Reg. 75191 (Oct. 30, 2023). https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence

266 Martins, L. (2025, February 4). Brazil's AI law faces uncertain future as big tech warms to Trump. Tech Policy Press. https://www.techpolicy.press/brazils-ai-law-faces-uncertain-future-as-big-tech-warms-to-trump/ ; Artificial Intelligence Act. (2025, January 9). South Korean AI Basic Law | Artificial Intelligence Act. https://artificialintelligenceact.com/south-korean-ai-basic-law/ ; Japan. (2021). **人工知能関連技術の研究開発及び活用の推進に関する法律.** In **法律**. https://www.cao.go.jp/houan/pdf/217/217anbun_2.pdf ; Department of Communications and Digital Technologies. (2024). Draft national artificial intelligence policy framework for South Africa [Draft]. https://www.dcdt.gov.za/sa-national-ai-policy-framework/file/338-sa-national-ai-policy-framework.html

267 United Nations Executive Office of the Secretary-General. (2023, May 24). A Global Digital Compact — an Open, Free and Secure Digital Future for All: Our Common Agenda Policy Brief 5. United Nations. https://www.un-ilibrary.org/content/papers/10.18356/27082245-28 ; Council of Europe. (2024). Framework convention on artificial intelligence and human rights, democracy and the rule of law (CETS No. 225)

SECTION 1
EQUALITY & NON-DISCRIMINATION

Equality and non-discrimination, as articulated by Article 1 of the UDHR and Article 2 of the ICCPR, state that "all human beings are born free and equal in dignity and rights". This means that the rights of every human being are universally equal, and that every person must be respected, protected, and provided with an opportunity to realize their human rights.

The presumption that technology is an inherently 'neutral' or 'objective' tool is misconstrued and short-sighted – they are built by humans, using data about humans. There is a wealth of evidence that technology, and especially data-driven tools such as AI and algorithms, can reinforce the wider structural power asymmetries they are embedded in. This means that technology may perpetuate and amplify structures of oppression and discrimination experienced by groups who are historically marginalised, such as racialised people, women and girls, low income communities, queer people, and disabled people.

AI and automated decision-making systems in the public sector and the proliferation of hatred and abuse on online platforms are particular areas of concern. Consequences of tech-facilitated discrimination can be severe, both on individuals and groups: racial discrimination arising from AI in policing has led to searches, questioning, arrests, detention, and harsher sentencing; while 'Tech-facilitated gender-based violence' (TfGBV) has led to stalking, threats, physical violence, and extreme mental distress.

To uphold the right to equality and non-discrimination in the digital age, the government should ensure that policy and legislation adheres to the following principles of digital rights:

1. Everyone is entitled to equal protection from all forms of discrimination posed by the design, development, and deployment of new and emerging digital technologies and online platforms, including, but not limited to, discrimination on the grounds of gender, race, social or economic class, sexuality, disability, religion, language, political opinion, national origin, marital status, and age.
2. The design, development and deployment of technology should be rooted first and foremost in impacts on communities' and individuals' human rights. When adopting new technologies into public sector applications, the state must undertake ex-ante and ex-post evaluation and monitoring to ensure that tech systems do not entrench existing societal bias and discrimination. States should abstain from developing and deploying technologies as catchall solutions without addressing underlying social problems.
3. States should support accessible and participatory policymaking processes and ongoing monitoring and evaluation of technology with communities most impacted by tech-facilitated discrimination and civil society groups, ensuring parity of participation between impacted groups and businesses and institutional interests in AI and tech governance conversations.

SECTION 2
PRIVACY & DATA PROTECTION

Privacy, as articulated in Article 12 of the UDHR and Article 17 of the ICCPR, is a foundational human right that enables us to live autonomously and in dignity. Following the digitisation of society, the right to data protection was introduced in international governance frameworks such as Convention 108 (1981) and the OECD guidelines (1980) to protect people against increasing privacy threats of widespread data exploitation by state, non-state and private actors.

Violations on the right to privacy and data protection are indicative of violations on other fundamental rights, such as non-discrimination; freedom of expression; freedom of thought, conscience, and religion; fair trial rights; freedom of movement; and others.

Potential threats to privacy and data protection stem from facial recognition, biometrics, digital ID, automated decision-making systems such as predictive policing, surveillance-based engagement methods on social media, and generative AI.

These uses, improperly deployed and inadequately restricted, can lead to undue surveillance, monitoring, abuse and manipulation, and unchecked power over people, especially marginalised groups.

To uphold the right to privacy in the digital age, the government should ensure that policy and legislation adheres to the following principles of digital rights:

1. Everyone has a right to protection from unnecessary and disproportionate uses of digital technology by state, non-state and private actors, that would undermine their ability to live in autonomy and dignity, or present unacceptable risks to their fundamental rights and freedoms.
2. Everyone has a right to know about, freely provide, or withdraw consent for, and challenge any measures to collect, aggregate, retain, and use their personal data.
3. Everyone has a right not to be subjected to profiling which could result in discriminatory impacts – whereby profiling is the processing of data about an individual's personality, behaviour, interests and habits, or proxies thereof to make predictions or decisions about them.
4. Everyone has a right to the protection of their personal information, as articulated in Convention 108. This includes the principles of:
    a. **Lawfulness, fairness and transparency:** Everyone has a right to lawful and fair processing of personal data, with clear and easily understandable information about that processing.

b. **Purpose limitation:** Personal data should only be collected for specified, explicit, and legitimate purposes.

c. **Data minimisation:** Processing of personal data must be adequate, relevant, and limited to what is necessary in relation to the purposes for which they are processed.

d. **Accuracy:** Data controllers must ensure that personal data is accurate and take every reasonable step to erase or rectify inaccurate data without delay.

e. **Storage limitation:** Personal data permitting identification of data subjects should not be kept longer than necessary, with established time limits.

f. **Integrity and confidentiality:** Personal data should be processed in a manner that ensures appropriate security, including protection against unauthorised access.

g. **Accountability:** Data controllers must take responsibility and demonstrate their compliance with the GDPR.

5. Everyone has a right to be forgotten in cases where the deletion of said data does not carry negative consequences for public interests. The right to be forgotten requires an entity holding their personal data to permanently destroy said data, including in cases where data would enable re-identification of the person.

6. Everyone has a right to unnecessary and disproportionate state, non-state, or corporate interference with online communications, preserved by access to private messaging platforms protected by end-to-end encryption.

SECTION 3
FREEDOM OF OPINION, EXPRESSION, INFORMATION & ASSEMBLY

The rights to freedom of opinion, expression, information & assembly, as articulated in Articles 19 and 20 of the UDHR and Articles 19 and 21 of the ICCPR, are crucial for living in an open, fair, and democratic society. They protect our rights to question the government and hold power accountable; attend protests; form social movements and organise politically; and communicate and connect with one another freely. The right also protects individuals' ability to seek, receive, and impart essential information.

The arrival of social media has brought new possibilities for exercising these rights and new opportunities for infringements, both of groups and individuals – through censorship, monitoring or surveillance, hate speech and online abuse. Additionally, Big Tech platform design such as recommender algorithms are having information online.

Balancing our rights on online platforms creates complex regulatory problems. However, we are seeing states responding with legislation that can be overly broad and vague, and is undermining people's rights to access and receive essential information.

The development and deployment of new surveillance technologies, particularly in criminal legal settings, is also known to have a 'chilling effect' on free expression, which is when people are led to modify their behaviour out of fear – such as remaining silent, avoiding certain places, or avoiding protests and demonstrations which are critical avenues for preserving democracy. To maintain human rights to freedom of opinion, expression, information & assembly in the digital age, the government should uphold the following principles of digital rights:

1. Everyone has the right to express themselves freely, including the right to speak, be heard, and participate in political, artistic, and social life, both offline and online – whereby the rights to freedom of expression and equality are mutually supporting and reinforcing human rights.[268]
2. Everyone has the right to seek, receive, and impart essential and rights-respecting and age-appropriate information freely, through any medium, including on the internet, without censorship or other arbitrary interference by state, non-state or private actors. This is particularly important for accessing vital information on mental or physical health, abuse, abortion, addiction, and others.

268 Human Rights Council, United Nations, & OHCHR. (2013). Annual report of the United Nations High Commissioner for Human Rights on the expert workshops on the prohibition of incitement to national, racial or religious hatred. In Annual Report of the United Nations High Commissioner for Human Rights (pp. 1–18) [Report]. https://www.article19.org/wp-content/uploads/2018/02/Rabat-Plan-of-Action-OFFICIAL-EN.pdf (Original work published 2013)

3. Everyone has the right to assemble and associate freely, including through and on the internet, for social, political, or cultural purposes, and without impacting their offline protest rights.
4. Everyone has the right to form their own opinion, without interference or manipulation. This is particularly important in relation to the influence of recommender algorithms on social media platforms in directing content to users.
5. Everyone has a right to protection from unnecessary and disproportionate uses of digital technology by state, non-state and private actors, that would undermine their ability to live in autonomy and dignity, or present unacceptable risks to their fundamental rights and freedoms.
6. States should abstain from passing legislation that creates opportunities for censorship to unduly infringe on human rights to free expression, access to essential information, and on- and offline assembly, in the present or future.
7. States should introduce legal frameworks to preserve freedom of expression on large online platforms that effectively function as online public spaces, often with real-world offline ramifications. These frameworks would need to both prevent undue suppression of online speech and assembly, and preserve people's rights to access quality information. Rights-respecting content moderation should include stipulations for due process; understandable rules and policies; cultural, linguistic, and contextual understanding; state involvement in content moderation; integrity and explainability; access to data; notices; and appeals.[269]

269 https://santaclaraprinciples.org/

SECTION 4
TRANSPARENCY, ACCOUNTABILITY & REDRESS

The right to effective remedy, as articulated by Article 8 of the UDHR and Article 2 of the ICCPR, ensures that individuals are able to seek redress for infringements on fundamental human rights. As such, the right is crucial for operationalising other fundamental rights.

This remains just as important for tech-induced infringements on human rights: people must be able to seek meaningful redress through judicial and non-judicial mechanisms in cases where technology has led to people to: suffer privacy infringements; be discriminated against or censored; be exposed to dis- or misinformation; be barred from accessing welfare, education, or employment; or experience online abuse.

However, there are many financial, technical, systemic, and regulatory issues that present barriers for individuals or groups seeking redress from tech harms. These barriers arise from fragmented liability and redress pathways, opaque technologies (notably AI) and organisations, and the cumulative and collective nature of many AI harms which can make them difficult to clearly evidence.

To uphold our right to an effective remedy in the digital age, the government should ensure that policy and legislation adheres to the following principles of digital rights:[270]

1. Everyone has a right to accessible and affordable judicial and non-judicial pathways for pursuing remedies against state, non-state, and private actors following tech-induced infringements on their rights. This should include pathways for both individuals and collectives to seek redress.[271]
2. Everyone has a right to access effective remedies against state, non-state and private actors following tech-induced infringements on their rights, which should constitute access to the following forms of remedy:[272] restitution, compensation, rehabilitation, satisfaction, guarantees of non-repetition, legal support, regulatory and legislative measures, and prevention.[273]

270 The principles in this section draw on the work of Yulu Pi and Maddie Proctor, whose paper on redress provides valuable analysis of the shortcomings of current redress mechanisms. See: 'Towards empowering AI governance with redress mechanisms'
271 On the need for collective data protection rights specifically, see Jeni Tennison's report 'Developing a Framework for Collective Data Rights': https://www.cigionline.org/publications/developing-a-framework-for-collective-data-rights/
272 United Nations Human Rights Office of the High Commissioner. (n.d.). Access to remedy and the technology sector: basic concepts and principles. https://www.ohchr.org/sites/default/files/access-to-remedy-concepts-and-principles.pdf
273 On the necessity of government-funded non-criminal redress schemes for victim-survivors to seek justice for intimate image abuse (IIA), see Glitch's position paper 'Beyond the Takedown: Non-criminal Redress for Intimate Image Abuse': https://glitchcharity.co.uk/our-work/non-criminal-redress

3. Everyone has a right to access information about the tech systems used by state, non-state and private actors, as a minimum requirement for enabling initiation of redress processes. These transparency measures should be evaluated according to people's ability to make informed decisions and better judgements relating to technology that impacts them.
    a. The information should include: the intended purpose of the system; how the system operates in practice; all data types and sources used by the system; what decisions or outcomes the system influences; and any internal reviews or evaluations.[274]
    b. Everyone has the right to be notified and made fully aware when they are interacting with an AI system or subject to an AI-driven decision.[275]
    c. Everyone has the right to clearly understand and identify the use of AI in any communicative or content output, ensuring that AI systems can be traced back to their origins.[276]
    d. Everyone has a right to clear and understandable resources and guidelines on accessing redress pathways that actively build people's capacity for seeking redress.
4. Everyone has a right to human review of decisions involving AI or algorithmic systems, with this option made available before the decision process is initiated.
5. States should introduce liability frameworks that clarify responsibility across the AI value chain for harms caused by AI – such as through mandating disclosure of evidence related to AI systems suspected of causing harm, or alleviating the burden of proof on claimants who face difficulties in demonstrating causation.[277]

274 Amnesty International UK. (2025). AUTOMATED RACISM: How police data and algorithms code discrimination into policing. https://www.amnesty.org.uk/files/2025-02/Automated%20Racism%20Report%20-%20Amnesty%20International%20UK%20-%202025.pdf
275 Public Law Project. (2023). Securing meaningful transparency of public sector use of AI. Comparative approaches across five jurisdictions. https://publiclawproject.org.uk/content/uploads/2024/10/Securing-meaningful-transparency-of-public-sector-AI.pdf
276 Gregory, S., & Llorente, R. V. (2023, October 31). Regulating transparency in Audiovisual Generative AI: How legislators can center human Rights. Tech Policy Press. https://www.techpolicy.press/regulating-transparency-in-audiovisual-generative-ai-how-legislators-can-center-human-rights/
277 Madiega, T. & European Parliamentary Research Service. (2023). Artificial intelligence liability directive. In EU Legislation in Progress (Report PE 739.342). https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/739342/EPRS_BRI(2023)739342_EN.pdf

SECTION 5
CORPORATE ACCOUNTABILITY & FAIR COMPETITION

The UN Guiding Principles on Business and Human Rights are a set of guidelines for states and companies to prevent, address, and remedy human rights abuses committed in business operations. The Guiding Principles require businesses to respect internationally recognised human rights, and states to prevent, investigate, punish and address human rights abuses by businesses.

In the digital age, a small number of technology companies have acquired enormous influence and power over the infrastructure and services shaping our online lives. Amazon, Microsoft, Google, Meta, and Apple – known as the Big Tech five – have acquired significant dominance over digital markets. Their market dominance introduces an opportunity for widespread threats to human rights by these actors' technologies and actions. For example, where surveillance-based business models threaten privacy rights; where engagement-driven methods amplify hateful content online threatening non-discrimination rights and manipulate access to reliable information; and where market domination limits market potential for rights-respecting digital alternatives. Additionally, the growing political influence of these companies may limit people's opportunities for accountability and redress as their lobbying activities are weakening initiatives for tech regulation.

The UK government should work to uphold and enforce the UN Business and Human Rights Guiding Principles in the digital age by ensuring that policy and legislation adheres to the following principles to help protect human rights in the digital age:[278]

1. States should recognise the critical role of existing regulatory frameworks in the UK for holding powerful technology companies accountable for human rights abuses – such as the GDPR, the Online Safety Act, competition and consumer protection laws, and the Digital Markets, Competition, and Consumers Act. Effective implementation of the GDPR is particularly important for addressing the human rights abuses of online platforms' surveillance-based business models.

278 The principles in this section draw on Amnesty Tech's reports 'Breaking up with Big Tech: a human rights-based argument for tackling Big Tech's market power' (2025) and 'Surveillance giants: How the business model of Google and Facebook threatens human rights' (2019), and the People Vs BigTech's 'Beyond Big Tech: A manifesto for a new digital economy' (2024). See: Amnesty International. (2026, January 6). Breaking up with Big Tech: A human rights-based argument for tackling Big Tech's market power. https://www.amnesty.org/en/documents/POL30/0226/2025/en/ ; https://www.amnesty.org/en/documents/pol30/1404/2019/en/ ; Amnesty International. (2021, June 1). Surveillance giants: How the business model of Google and Facebook threatens human rights. https://www.amnesty.org/en/documents/pol30/1404/2019/en/ ; and Abdul-Rahim, R. (2025, October 22). Beyond Big Tech: A manifesto for a new digital economy. People Vs. Big Tech. https://peoplevsbig.tech/beyond-big-tech-a-manifesto-for-a-new-digital-economy/

2. Technology companies should be expected to carry out human rights due diligence evaluations on all aspects of their operations to identify and address human rights impacts related to their global operations, and implement effective enforcement mechanisms to ensure companies are held legally accountable for human rights harms.
3. Everyone has the right to effectively and freely choose, switch, and transfer their information between different platforms and services, without losing networks or content, and based on objective, transparent, easily accessible, and reliable information.
4. States should introduce measures to counter the outsized lobbying power of Big Tech to protect against infringements on human rights. This could involve conflict-of-interest rules for public officials moving from lobbying roles at technology companies and transparency requirements on lobbying activity.

SECTION 6
ADEQUATE STANDARD OF LIVING

Article 25 of the UDHR and Article 11 of the ICESCR ensures that everyone has access to an adequate standard of living for themselves and their families, which includes food, clothing, housing, healthcare, and social security. The right obliges the state to protect and preserve basic needs, and states generally deliver on this obligation through the provision of essential public services such as healthcare, housing, and welfare, and is closely connected to other economic, social, and cultural rights such as the right to health and social security.

The digital age is bringing profound changes to the functioning of essential public services. As part of the government's wider drive for AI adoption in the public sector, many public sector authorities are making significant investments to digitise their services.

While the benefits for efficiency in particular are promising, the wave of new tech adoption could also threaten human rights to privacy, non-discrimination, and economic, social and cultural rights. For example, the use of personal data and inferred characteristics to make 'risk predictions' and key welfare decisions could unfairly prevent people from accessing the services they need. We have already seen cases of people being inappropriately denied life-saving surgery, subject to fraud investigation,[279] and having social benefits cut off.[280] In addition, automation is bringing changes to the labor market, leading to growing fears of job displacement, which would leave people without income, and in need of essential social security support.[281]

To uphold a right to an adequate standard of living in the digital age, the government should ensure that policy and legislation for the automation of essential public services adheres to the following principles of digital rights:

1. Everybody has the right to access all essential public services through equal non-digital means for those who cannot or do not want to apply online.
2. Everyone has a right to meaningful transparency measures around tech systems used by essential public services that are evaluated according to people's ability to seek redress.
3. Everyone has a right to opt-out of decisions involving AI or algorithmic systems in essential public services.

279 Booth, R. (2025, April 2). DWP algorithm wrongly flags 200,000 people for possible fraud and error. The Guardian. https://www.theguardian.com/society/article/2024/jun/23/dwp-algorithm-wrongly-flags-200000-people-possible-fraud-error
280 Heikkilä, M. (2022, April 13). Dutch scandal serves as a warning for Europe over risks of using algorithms. POLITICO. https://www.politico.eu/article/dutch-scandal-serves-as-a-warning-for-europe-over-risks-of-using-algorithms/
281 Frazier, K. (2026, January 24). We need a new kind of insurance for AI job loss. AI Frontiers. https://ai-frontiers.org/articles/ai-displacement-insurance

4. Everyone has a right to human review of decisions involving AI or algorithmic systems in essential public services.
5. The deployment of technologies in public services should prioritise impacts on communities' and individuals' human rights, oriented toward solving the specific challenges limiting the quality of services delivered to the public. These are often complex systemic challenges underpinned by a myriad of social and structure problems. Where new technologies are offered as a solution but do not address the underlying issues, their application will have limited impact and can risk further entrenching the existing biases and social and structural challenges.

SECTION 7
FREEDOM OF MOVEMENT & ASYLUM

The rights to freedom of movement and asylum, as articulated by Articles 13 and 14 of the UDHR and Article 12 of the ICCPR, enshrine the ability for each person to live a dignified life that is safe from persecution. These rights are crucial to give people an opportunity to pursue a better life, free from unfair punishment, torture and marginalisation. As global conflict, climate change and economic insecurity continue to drive people around the world to seek stability and security, these rights remain as pertinent as they were during their initial introduction, yet they are increasingly threatened by an intensifying backdrop of hostility towards migrants.

Against this backdrop, the implementation of new digital technologies by states for border and migration control has the potential to further undermine rights to freedom of movement and asylum if proper restrictions and safeguards are not put in place – in particular, exclusionary 'digital by default' systems, intimidatory surveillance, remote biometric identification (RBI), and discriminatory automated decision-making and risk assessment technologies. These technologies can also infringe on people's rights to privacy and non-discrimination, as well as key economic, social and cultural rights such as the right to work, health, social security, and adequate standard of living – as per other sections in this declaration.

To uphold the right to freedom of movement and asylum in the digital age, the government should ensure that policy and legislation adheres to the following principles of digital rights:[282]

1. Everyone has a right to protection in migration contexts from unnecessary and disproportionate uses of digital technology by state, non-state and private actors, that would undermine their ability to move or settle freely, in autonomy and dignity, or present unacceptable risks to their fundamental rights and freedoms.
2. The design, development and deployment of technology in migration contexts should be rooted first and foremost in impacts on communities' and individuals' human rights, instead of justifications for efficiency, innovation, cost-cutting or other political agendas.
3. Everybody has the right to access all systems and processes in migration contexts through equal non-digital means. There must be alternative application options for those who cannot or do not want to apply online.
4. Everyone has a right to meaningful transparency measures around technologies in migration settings that are evaluated according to people's ability to seek redress.
5. Everyone subjected to technologies in migration settings has the right to know about, freely provide or withdraw consent for, and challenge any measures to collect, aggregate, retain, and use their personal data·

282 For more granular recommendations on regulating technology in migration contexts, see Amnesty International and the #ProtectNotSurveil coalition's 'Advocacy Briefing for Defending the Rights of Refugees, Asylum Seekers, and Migrants in The Digital Age': https://www.amnesty.org/en/documents/pol30/0290/2025/en/?trk=feed_main-feed-card_feed-article-content

SECTION 8
FAIR & JUST WORKING CONDITIONS

The right to work, as articulated by Article 23 of the UDHR and Articles 6, 7, and 8 of the ICESCR, entitle everyone to key workplace rights, including: just and favourable conditions of work; protection against unemployment; fair and equal pay, and equal work; and the right to form and join trade unions.

Workplace technologies can bring benefits to workers if introduced correctly, but there are many ways they can threaten fundamental rights and reinforce existing labor inequalities.

Threats to workers' digital rights include: privacy concerns associated with new methods for surveillance and monitoring as well as chilling effects on freedom of expression and association; algorithmic and AI-induced discrimination; growing job displacement, precarity, and deskilling; increased stress and safety concerns; and reduced autonomy, agency and wellbeing at work.

Additionally, the emergence of on-demand platforms such as Uber, Deliveroo, and JustEat, has given rise to a new workforce of gig economy workers who are often most impacted by automation in the workplace – such as through automated firing and hiring algorithms, or dynamic pay systems – but are excluded from key rights under UK employment law. Additional employment protections are urgently needed to ensure that the rights of all workers are fulfilled in the digital age.

To uphold the right to work in the digital age, the government should ensure that policy and legislation adheres to the following principles of digital rights:

1. Every worker across the digital value chain has the right to just and favourable working conditions as defined in the UDHR and ICESCR, both online and offline.
2. Every worker has the right to a single employment status and entitlement to all protections under UK employment law, irrespective of employment classification.
3. Every worker has the right to participate in, be consulted, and negotiate on the design, development, deployment and ongoing monitoring of AI and algorithmic systems in the workplace, with universal rights for collective bargaining.
4. Every worker has a right to gain a fair share and agency over any gains from technological progress, and to protection from any downsides.
5. Every worker has the right to switch off or disconnect from digital working environments.
6. Every worker whose job is threatened by displacement through AI or automation has the right to retrain, and access social security in cases of unemployment because of job displacement.

7. Every worker has the right to whistle-blower protections to support public accountability against developers and deployers of AI technologies.
8. States should introduce redlines around workplace uses of technology that present unacceptable risks for workers rights, considering, for example, certain applications of affective cognition technology, dynamic pay systems, and automated firing and hiring systems.
9. States should introduce robust legal frameworks and safeguards, such as those set out in the Trades Union Congress Artificial Intelligence (Regulation and Employment Rights) Bill,[283] to ensure that workers are not subjected to workplace uses of AI, algorithms, or digital technology that are incompatible with their fundamental rights, and to ensure that employers are required to enable the realisation of all rights and principles in this declaration.
10. States should equally and meaningfully involve trade unions as representative bodies of workers in policymaking processes around tech regulation, to ensure that workers have parity of participation with businesses and institutional interests in AI and tech governance conversations.

283 Trades Union Congress. (2024). The AI Bill Project. https://www.tuc.org.uk/research-analysis/reports/ai-bill-project

SECTION 9
CHILDREN'S RIGHTS

The UN Convention on the Rights of the Child (UNCRC) outlines 54 articles covering all aspects of a child's life and apply universally to every child. The UNCRC General Comment No. 25 (2021) on children's rights in the digital environment states: "The rights of every child must be respected, protected and fulfilled in the digital environment."

The extensive influence of technology on children's lives has consequences for many rights in the UNCRC – children are impacted by all the other rights in this declaration which apply to adults, whether privacy, non-discrimination, and freedom of expression. However, it is also important to single out children's digital rights because of the specific vulnerabilities of children in relation to digital technology, and the importance for the future well-being of society as a whole to give special consideration to children's rights. Additionally, given the increasing adoption of EdTech tools, and the proliferation of child sexual abuse material (CSAM) on online platforms, special attention is needed for children's rights to education and protection from exploitation in the digital age.

To ensure the rights of children are respected, protected, and fulfilled in a rapidly evolving digital age, the government should ensure that policy and legislation adheres to the following principles of digital rights:

1. States should fulfill their obligations regarding the best interests of the child in relation to design, development, and deployment of digital products and services that are likely to impact children, whether by state, non-state or private companies. Best interests of the child means respecting, protecting, and fulfilling all rights in the UNCRC in concert with one another, and realising all rights and principles within this declaration for children.[284]
2. Additionally, of particular relevance for children is the right to education in the digital age. The digitalisation of education should not impact children's right to education, which entails every child's right to free, quality, public education without increasing inequalities or leading to violations of other rights within education, particularly the right to privacy. Digital skills education is particularly important for enabling children's digital inclusion and ability to exercise their right to education.
3. States shall take all appropriate legislative, administrative, social and educational measures to protect the child from all forms of physical or mental violence, injury or abuse, neglect or negligent treatment, maltreatment or exploitation, including sexual abuse – both online and offline.

284 Livingstone, S. et al. (2024). The best interests of the child in the digital environment. https://www.digital-futures-for-children.net/digitalfutures-assets/digitalfutures-documents/Best-Interests-of-the-Child-FINAL.pdf

4. Any legislation passed to protect the safety of children in relation to digital technologies should meet the best interests of the child, where the best interests of the child means fulfilling all rights in the UNCRC in concert. This is particularly important for legislation regulating online platforms that carries risks for children's right to free expression, access to information, and assembly, and their right to privacy. Every child has the right to access free, good quality, age appropriate and non-discriminatory information, especially mental health and sexual and reproductive health information.[285] Equally, every child has a right to access good quality information through non-digital alternatives for those who cannot or do not want to access essential information online.
5. Every child has the right to access digital products that are safe by design, and that this design strategy is met throughout the entire lifecycle of new technologies – from design, to deployment, to retirement.
6. Every child has the right to media, information and AI literacy, and should be offered adequate support to fully access and exercise this right, whereby media, information, and AI literacy is the knowledge and skills to be able to participate in the digital media environment as creative media producers, active citizens and critical information consumers in an age appropriate way. This should include developing a critical understanding of the technologies and systems that influence the online environment.

285 United Nations Human Rights Office of the High Commissioner. (2021). General comment No. 25 on children's rights in relation to the digital environment. https://www.ohchr.org/en/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-relation

SECTION 10
DIGITAL INCLUSION

Digital inclusion (e.g. access to internet connection and adequate digital tools and skills education) is not an established human right, but it is increasingly essential for people to fully realise their fundamental human rights in our modern digital age. Digital inclusion should therefore be committed to in a declaration on digital rights as an essential enabler of existing human rights.

Basic functionalities of day-to-day life are increasingly mediated or accessed through digital technology: whether making or receiving payments, booking medical appointments, applying for jobs, accessing benefits, or communicating with friends and family. Accordingly, digital inclusion, the equitable and safe access to digital technologies and digitally mediated services and opportunities, is becoming essential for realising a variety of rights, including rights to social security, work, education, and health.

However, there are many people across the UK who face significant barriers to digital inclusion. High costs of digital devices and internet connection prevent millions of people from getting and staying online.[286] Inadequate digital skills combined with exclusionary technology design hinders individuals from using digital technology and services effectively, or to challenge the technology companies.[287]

Digital exclusion disproportionately affects certain groups in the UK.[288] These are: low-income households, older people, disabled people, people experiencing unemployment and seeking work, young people (including those not in education, employment or training), individuals with low levels of English literacy, refugees and asylum seekers, homeless individuals who have been displayed due to domestic violence, human trafficking, or modern slavery, and people living in rural communities.

To uphold a range of fundamental human rights including non-discrimination, social security, education, and health, the government should introduce the right to digital inclusion as indivisible from other fundamental human rights which depend on it.

The government should ensure that policy and legislation adheres to the following principles of digital rights:

286 Citizens Advice. (2023, May 18). One million lose broadband access as cost-of-living crisis bites. https://www.citizensadvice.org.uk/about-us/media-centre/press-releases/one-million-lose-broadband-access-as-cost-of-living-crisis-bites/
287 Carmi, E., & Yates, S. (2023, May 18). Civic Participation in the Datafied Societyl Data Citizenship: Data Literacies to challenge power imbalance between society and "Big Tech." https://ijoc.org/index.php/ijoc/article/view/18823
288 Digital Inclusion Action Plan: summary of responses (published 17 July 2025). (2025, July 17). GOV.UK. https://www.gov.uk/government/calls-for-evidence/digital-inclusion-action-plan/outcome/digital-inclusion-action-plan-summary-of-responses-published-17-july-2025

1. Digital inclusion is instrumental for realising human rights in our digital age and should therefore be actively implemented in policy and legislation as a critical human rights enabler. Doing so might involve:
    a. Enshrining key protections and provisions for digital inclusion, such as by transitioning the current Digital Inclusion Action Plan into a long-term, enforceable, and legally-binding Digital Inclusion Act.
    b. Unify digital inclusion metrics and build the digital inclusion evidence-base, to improve understanding of digital inclusion, and how this impacts people's engagement with support and outcomes.
2. Everybody has the right to internet connectivity, through affordable, reliable, high-quality internet with regulated broadband pricing and automatically applied social tariffs for eligible users. The government might support this right by enacting policies prohibiting the arbitrary and deliberate slowing and/or cutting off of the internet, including public order or national security.
3. Everyone has the right to access digital devices that are essential to exercising their other rights should they wish to use digital devices. To help support this right the government might, for example, mandate the refurbishment and redistribution of surplus or outdated digital devices from public bodies to people without access to devices, as opposed to resale, waste or recycling disposal.
4. Everyone has a right to request non-digital options for accessing public services for those who cannot or do not wish to to use digital services. This is to ensure autonomy and freedom of choice. Non-digital options should also be easy to access and use and not carry additional burdens.
5. Everyone has the right to free and non-corporate digital literacy education, including accessible, local community-based support. Such training should be made available throughout different life stages and in a diversity of settings, recognising that peoples' access, skills, confidence and motivation can evolve in different contexts just as technology evolves. Support should also cover technological changes and device failures.

SECTION 11
HEALTHY ENVIRONMENT

In 2022, the UN passed a resolution declaring that everyone on the planet has the right to a healthy environment. Article 37 of the EU Charter of Fundamental Rights establishes a legally-binding human right to a favourable environment. A similar right has been proposed by the Scottish Government as an amendment to their Human Rights Bill; it stipulates that the right to a healthy environment includes "clean air; safe climate; healthy ecosystems and biodiversity; access to safe and sufficient water; healthy and sustainably produced food; and non-toxic environments in which to live, work and play."[289]

There is growing awareness and understanding of the detrimental environmental impacts caused by technology across the whole digital supply chain – from mining processes to extract critical minerals for hardware such as lithium, cobalt, or copper; to the energy intensive data centres housing AI's compute capacity; to the enormous quantities of dumped digital waste. These all have detrimental environmental impacts on local communities, both in the UK and internationally, undermining the fundamental right to a healthy environment.

To uphold the right to a healthy environment, the UK government should ensure that policy and legislation adheres to the following principles:

1. States should take steps to promote a 'circular economy' around digital technologies as a sustainable business model that eliminates waste by keeping materials in use for as long as possible. This should involve:
    a. Mandating sustainable and responsible corporate behaviour and business models throughout the global supply chains of digital products and services.
    b. Promoting digital product services and strategies that prioritise and promote repairs, recycling and re-use.
    c. Introducing bespoke resource disclosure frameworks, consumption reduction regulations and targets, and environmental and social impact assessments for the technology sector to report, reduce, and responsibly manage their carbon- and material-based consumption.
    d. Prohibiting planned obsolescence, or the deliberate reduction of a product's lifecycle by developers.

289 The Scottish Government. (2025, July 3). Human Rights Bill for Scotland: discussion paper. https://www.gov.scot/publications/human-rights-bill-scotland-discussion-paper/pages/7/

2. States should engage in harmonised global collaboration on emission reduction targets for the technology supply chain to minimise negative environmental impacts of digital technology.
3. Everyone has the right to participate in, be consulted, and negotiate on the planning permission, design, and construction of major tech infrastructure projects that could have potential implications on the health, protection, and sustainability of their local environment.

SECTION 12
LIFE, LIBERTY & SECURITY OF PERSON

Article 6 of the ICCPR states that "Every human being has the inherent right to life. This right shall be protected by law. No one shall be arbitrarily deprived of his life.". The right to life means that nobody, including governments, can try to end your life, and that governments should take appropriate measures to safeguard life by making laws to protect you.

Debates around the right to life and what constitutes an infringement commonly arise with respect to military and law enforcement operations. The right to life requires that the use of force is necessary to achieve a legitimate aim and applied in a proportionate manner. The right also requires that lethal force may only be used as a last resort to protect human life.[290]

Increasingly, governments are deploying AI and data-driven technologies in military and law enforcement domains,[291] ranging from decision support software, target identification, drone and satellite analysis of conflict zones, prediction of resource needs, and attack by lethal autonomous weapons, with implications for right to life and security of person. Recent examples include the Israeli military's AI systems used in Gaza such as "The Gospel" and "Lavender" for generating bombing targets – both buildings and individuals.[292]

These powerful and unchecked technologies create pressing dangers for international peace and security, escalating arms proliferation, and accelerating the speed and scale of war. The dangers of military uses of AI, and especially autonomous weapons systems, raise serious threats to fundamental obligations and principles of international human rights law,[293] including the right to life, as well as human dignity, privacy, remedy, and non-discrimination – as well as profound concerns for international humanitarian law.

Diplomatic efforts to govern military AI are intensifying.[294] The REAIM summits hosted by the Netherlands, South Korea and Spain, have convened over 90 countries to advance global dialogue on international norms for AI in the military domain.[295] In 2024, the UN General Assembly adopted resolution 79/239 on "Artificial intelligence in the military domain and its implications for international peace and security".[296] However, there is currently no international governance framework for regulating AI in the military domain. Global collaboration is urgently needed to establish robust international governance frameworks on AI in the military domain.

290 https://www.hrw.org/report/2025/04/28/hazard-human-rights/autonomous-weapons-systems-and-digital-decision-making ; https://international-review.icrc.org/sites/default/files/irrc_864_8_0.pdf
291 Privacy International. (2025). What is militarisation of tech? https://privacyinternational.org/long-read/5668/what-militarisation-tech
292 Questions and Answers: Israeli military's use of digital tools in Gaza. (2024, October 3). Human Rights Watch. https://www.hrw.org/news/2024/09/10/questions-and-answers-israeli-militarys-use-digital-tools-gaza#_What_are_some
293 Docherty, B. (2025). A hazard to human rights. In Human Rights Watch. https://www.hrw.org/report/2025/04/28/hazard-human-rights/autonomous-weapons-systems-and-digital-decision-making ; Artificial intelligence in the military domain: ICRC submits recommendations to UN Secretary-General. (2025, April 17). International Committee of the Red Cross. https://www.icrc.org/en/article/artificial-intelligence-military-domain-icrc-submits-recommendations-un-secretary-general
294 Csernatoni, R. (2024, July 17). Governing military AI amid a geopolitical minefield. Carnegie Endowment for International Peace. https://carnegieendowment.org/research/2024/07/governing-military-ai-amid-a-geopolitical-minefield?lang=en
295 Ministry of Foreign Affairs, Republic of Korea. REAIM Regional Consultations held in Asia-Pacific ViewlPress Releases | Ministry of Foreign Affairs, Republic of Korea. https://www.mofa.go.kr/eng/brd/m_5676/view.do?seq=322853
296 United Nations Office for Disarmament Affairs. (2024). Artificial intelligence in the military domain. https://disarmament.unoda.org/en/our-work/emerging-challenges/artificial-intelligence-military-domain

In line with recommendations from the International Committee of the Red Cross (ICRC) and the International Committee for Robot Arms Control (ICRAC), the UK government should adhere to the following principles of digital rights protect life and personal security in an era of AI-enabled weapons:

1. Everyone has an equal right to life as currently interpreted by established human rights frameworks, including in contexts of law enforcement and armed conflict. These rights should not be lessened or weakened by the introduction of new technologies.
2. Where necessary, regulatory steps and prohibitions should be implemented to preserve the rights to life, liberty and security of person in relation to AI applications in the military domain. These interventions may include the following recommendations from the International Committee of the Red Cross (ICRC):
    a. Prohibitions against unpredictable autonomous weapon systems that do not allow a human user to understand, explain or predict the system's functioning and effects.
    b. Prohibitions against autonomous weapon systems designed or used to target humans directly. This is required because of the significant risk of IHL violations and the unacceptability of anti-personnel autonomous weapons from an ethical perspective.
    c. Strictly regulate meaningful human control in all other considerations related to the development and use of AI in military applications. It is essential that human control and judgement are preserved in decisions that pose risks to the life and dignity of people affected by armed conflict.
3. The state should support the establishment and operationalisation of an international legally-binding treaty regulating the military application of AI. As a starting point, such a treaty should leverage ongoing efforts such as REAIM and the UN processes[297] and aim to complement existing international frameworks for humanitarian and human rights law.

297 Csernatoni, R. (2024, July 17). Governing military AI amid a geopolitical minefield. Carnegie Endowment for International Peace. https://carnegieendowment.org/research/2024/07/governing-military-ai-amid-a-geopolitical-minefield?lang=en

Licence to publish

Demos – Licence to Publish

The work (as defined below) is provided under the terms of this licence ('licence'). The work is protected by copyright and/or other applicable law. Any use of the work other than as authorized under this licence is prohibited. By exercising any rights to the work provided here, you accept and agree to be bound by the terms of this licence. Demos grants you the rights contained here in consideration of your acceptance of such terms and conditions.

1 Definitions

a 'Collective Work' means a work, such as a periodical issue, anthology or encyclopedia, in which the Work in its entirety in unmodified form, along with a number of other contributions, constituting separate and independent works in themselves, are assembled into a collective whole. A work that constitutes a Collective Work will not be considered a Derivative Work (as defined below) for the purposes of this Licence.

b 'Derivative Work' means a work based upon the Work or upon the Work and other pre-existing works, such as a musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which the Work may be recast, transformed, or adapted, except that a work that constitutes a Collective Work or a translation from English into another language will not be considered a Derivative Work for the purpose of this Licence.

c 'Licensor' means the individual or entity that offers the Work under the terms of this Licence.

d 'Original Author' means the individual or entity who created the Work.

e 'Work' means the copyrightable work of authorship offered under the terms of this Licence.

f 'You' means an individual or entity exercising rights under this Licence who has not previously violated the terms of this Licence with respect to the Work, or who has received express permission from Demos to exercise rights under this Licence despite a previous violation.

2 Fair Use Rights

Nothing in this licence is intended to reduce, limit, or restrict any rights arising from fair use, first sale or other limitations on the exclusive rights of the copyright owner under copyright law or other applicable laws.

3 Licence Grant

Subject to the terms and conditions of this Licence, Licensor hereby grants You a worldwide, royalty-free, non-exclusive, perpetual (for the duration of the applicable copyright) licence to exercise the rights in the Work as stated below:

a to reproduce the Work, to incorporate the Work into one or more Collective Works, and to reproduce the Work as incorporated in the Collective Works;

b to distribute copies or phono-records of, display publicly, perform publicly, and perform publicly by means of a digital audio transmission the Work including as incorporated in Collective Works; The above rights may be exercised in all media and formats whether now known or hereafter devised. The above rights include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. All rights not expressly granted by Licensor are hereby reserved.

4 Restrictions

The licence granted in Section 3 above is expressly made subject to and limited by the following restrictions:

a You may distribute, publicly display, publicly perform, or publicly digitally perform the Work only under the terms of this Licence, and You must include a copy of, or the Uniform Resource Identifier for, this Licence with every copy or phono-record of the Work You distribute, publicly display, publicly perform, or publicly digitally perform. You may not offer or impose any terms on the Work that alter or restrict the terms of this Licence or the recipients' exercise of the rights granted hereunder. You may not sublicence the Work. You must keep intact all notices that refer to this Licence and to the disclaimer of warranties. You may not distribute, publicly display, publicly perform, or publicly digitally perform the Work with any technological measures that control access or use of the Work in a manner inconsistent with the terms of this Licence Agreement. The above applies to the Work as incorporated in a Collective Work, but this does not require the Collective Work apart from the Work itself to be made subject to the terms of this Licence. If You create a Collective Work, upon notice from any Licensor You must, to the extent practicable, remove from the Collective Work any reference to such Licensor or the Original Author, as requested.

b You may not exercise any of the rights granted to You in Section 3 above in any manner that is primarily intended

for or directed toward commercial advantage or private monetary compensation. The exchange of the Work for other copyrighted works by means of digital file sharing or otherwise shall not be considered to be intended for or directed toward commercial advantage or private monetary compensation, provided there is no payment of any monetary compensation in connection with the exchange of copyrighted works.

c If you distribute, publicly display, publicly perform, or publicly digitally perform the Work or any Collective Works, you must keep intact all copyright notices for the Work and give the Original Author credit reasonable to the medium or means You are utilizing by conveying the name (or pseudonym if applicable) of the Original Author if supplied; the title of the Work if supplied. Such credit may be implemented in any reasonable manner; provided, however, that in the case of a Collective Work, at a minimum such credit will appear where any other comparable authorship credit appears and in a manner at least as prominent as such other comparable authorship credit.

5 Representations, Warranties and Disclaimer

a By offering the Work for public release under this Licence, Licensor represents and warrants that, to the best of Licensor's knowledge after reasonable inquiry:

i Licensor has secured all rights in the Work necessary to grant the licence rights hereunder and to permit the lawful exercise of the rights granted hereunder without You having any obligation to pay any royalties, compulsory licence fees, residuals or any other payments;

ii The Work does not infringe the copyright, trademark, publicity rights, common law rights or any other right of any third party or constitute defamation, invasion of privacy or other tortious injury to any third party.

b Except as expressly stated in this licence or otherwise agreed in writing or required by applicable law, the work is licenced on an 'as is' basis, without warranties of any kind, either express or implied including, without limitation, any warranties regarding the contents or accuracy of the work.

6 Limitation on Liability

Except to the extent required by applicable law, and except for damages arising from liability to a third party resulting from breach of the warranties in section 5, in no event will licensor be liable to you on any legal theory for any special, incidental, consequential, punitive or exemplary damages arising out of this licence or the use of the work, even if licensor has been advised of the possibility of such damages.

7 Termination

a This Licence and the rights granted hereunder will terminate automatically upon any breach by You of the terms of this Licence. Individuals or entities who have received Collective Works from You under this Licence, however, will not have their licences terminated provided such individuals or entities remain in full compliance with those licences. Sections 1, 2, 5, 6, 7, and 8 will survive any termination of this Licence.

b Subject to the above terms and conditions, the licence granted here is perpetual (for the duration of the applicable copyright in the Work). Notwithstanding the above, Licensor reserves the right to release the Work under different licence terms or to stop distributing the Work at any time; provided, however that any such election will not serve to withdraw this Licence (or any other licence that has been, or is required to be, granted under the terms of this Licence), and this Licence will continue in full force and effect unless terminated as stated above.

8 Miscellaneous

a Each time You distribute or publicly digitally perform the Work or a Collective Work, Demos offers to the recipient a licence to the Work on the same terms and conditions as the licence granted to You under this Licence.

b If any provision of this Licence is invalid or unenforceable under applicable law, it shall not affect the validity or enforceability of the remainder of the terms of this Licence, and without further action by the parties to this agreement, such provision shall be reformed to the minimum extent necessary to make such provision valid and enforceable.

c No term or provision of this Licence shall be deemed waived and no breach consented to unless such waiver or consent shall be in writing and signed by the party to be charged with such waiver or consent.

d This Licence constitutes the entire agreement between the parties with respect to the Work licenced here. There are no understandings, agreements or representations with respect to the Work not specified here. Licensor shall not be bound by any additional provisions that may appear in any communication from You. This Licence may not be modified without the mutual written agreement of Demos and You.

DEMOS

Demos is a champion of people, ideas and democracy. We bring people together. We bridge divides. We listen and we understand. We are practical about the problems we face, but endlessly optimistic and ambitious about our capacity, together, to overcome them.

At a crossroads in Britain's history, we need ideas for renewal, reconnection and the restoration of hope. Challenges from populism to climate change remain unsolved, and a technological revolution dawns, but the centre of politics has been intellectually paralysed. Demos will change that. We can counter the impossible promises of the political extremes, and challenge despair – by bringing to life an aspirational narrative about the future of Britain that is rooted in the hopes and ambitions of people from across our country.

Demos is an independent, educational charity, registered in England and Wales. (Charity Registration no. 1042046)

Find out more at **www.demos.co.uk**

DEMOS

PUBLISHED BY DEMOS FEBRUARY 2026

15 WHITEHALL, LONDON, SW1A 2DD
T: 020 3878 3955
HELLO@DEMOS.CO.UK
WWW.DEMOS.CO.UK