Policy implications of artificial intelligence (AI)

Overview

- Artificial intelligence (AI) is developing at a rapid pace and can be found throughout society in a growing range of everyday applications and decision-making. There are implications for security, privacy, transparency, liability, labour rights, intellectual property and disinformation. It presents some risks and benefits to democracy more widely.
- There is no dedicated AI legislation in the UK. Existing legislation restricts how AI can be used in practice, such as in relation to data protection, equality and human rights, and intellectual property.
- In March 2023, the UK Government announced a 'pro-innovation' approach to AI regulation, which largely regulates AI via existing laws enforced by existing regulators. It outlined cross-sectoral principles, such as safety, security, robustness, transparency, fairness, accountability, contestability, and redress, for existing regulators to consider. The approach applies to the whole of the UK, although some policy areas are devolved.
- The Government has brought forward legislation and regulatory action on automated vehicles and data protection and digital information.
- Some stakeholders have indicated that additional legislation and action may be required, including mandatory impact assessments, bans on certain AI applications, and a right for human intervention to challenge AI decision-making. There are concerns that regulators are not currently equipped with the staffing, expertise or funding to regulate AI.

Background

Artificial intelligence (AI) can be found in a wide variety of everyday applications, such as in deciding what users see on social media[1,2] digital personal assistants,[3] and recognising patterns in images for medical diagnosis.[4–7]

AI technology is described in Artificial intelligence: An explainer. This briefing focuses on policy aspects.

AI systems are increasingly being used in the public and private sector for decision-making.[8] Some current and potential future examples of automated decision-making include:

- assisting with managing workers, such as allocating work, monitoring performance, and determining pay[9]

- aiding with local and national public sector decisions, such as social housing allocation,[10] benefit claims, and other issues[10–15]

- self-driving cars[*16,17]

In the past few years, there have been significant advancements in AI capabilities. Single AI systems can now perform a wide range of tasks (PB 57). Generative AI (see Table for definitions) can generate realistic text, images, audio, and video.[18–20]

There have been significant public policy developments relating to AI in recent years.

In 2023, the Government published a white paper outlining a 'pro-innovation approach to AI',[21] along with announcements including £900m for an 'exascale' [†] supercomputer.[22] The Prime Minister hosted a global 'AI Safety Summit' in November 2023.[23] It resulted in a declaration on AI safety signed by 28 countries,[24] and the establishment of an AI Safety Institute[25] that aims to build public sector capacity to research AI safety.

The rapid advancement of AI capabilities means that policy implications are continuously shifting. Some stakeholders have called for further Government action.

---

[*] The Kings Speech announced an Automated Vehicles Bill that seeks to introduce new legal frameworks to support safe commercial developments of self-driving vehicles.[16]

[†] An exascale supercomputer is a computer that can perform 10^{18} (a quintillion) operations per second.

Table: Definitions of AI and related concepts

Definitions are not universally agreed, are continuously evolving, and are linked.

Term	Definition
Artificial intelligence (AI)	The UK Government's 2023 policy paper on 'A pro-innovation approach to AI regulation' defined AI, AI systems or AI technologies as "products and services that are 'adaptable' and 'autonomous'." The adaptability of AI refers to AI systems, after being trained, often developing the ability to perform new ways of finding patterns and connections in data that are not directly envisioned by their human programmers. The autonomy of AI refers to some AI systems that can make decisions without the intent or ongoing control of a human ([PB 57](#)).
Generative AI	The Alan Turing Institute defines generative AI as an "artificial intelligence system that generates text, images, audio, video or other media in response to user prompts."[26] Generative AI applications include chatbots, such as OpenAI's ChatGPT, photo and video filters, and virtual assistants ([PB 57](#)).
Automated decision-making	A term that the Office for AI* uses to refer to "both solely automated decisions (no human judgement involved) and automated assisted decision-making (assisting human judgement)."[27]
Training datasets	The set of data used to train an AI system, which often requires labelling (such as captioning pictures) to explain what the data means.
Algorithm	A set of instructions used to perform tasks (such as calculations and data analysis) usually using a computer or another smart device ([PB 57](#)).

* The Office for AI is an office within the Department for Science, Innovation, and Technology.

Benefits and risks

Job creation and displacement

In 2022, a report commissioned for the Department for Science, Innovation & Technology stated that investment in the UK AI sector grew five-fold between 2019 and 2021, and that it brought in £10.6 billion in revenue and employed over 50,000 people.[28]

OECD analysis published in July 2023 found that AI was changing the nature of work by assisting workers and reducing the time spent on mundane tasks, rather than causing job losses.[29]

However, some emerging academic research indicates that developments in generative AI may be linked to a loss in the quantity and earnings of white-collar jobs.[30,31]

There is potential for new jobs across all sectors to be created with improved productivity and global economic growth (PB 57).[32] However, some reports state that certain jobs, such as clerical work,[33,34] could become redundant.

Stakeholders have raised concerns that AI developments may disproportionately affect disadvantaged groups. For example, the majority of clerical work is carried out by women.[34–37]

A 2021 report commissioned by the former Department for Business, Energy and Industrial Strategy highlighted a regional disparity in the net employment impacts of AI, with London and the South East benefitting from net job gains more than Northern England and the Midlands.[34]

Some academics and think tanks,[38] and technology trade associations,[39] have said that the Government should help workers retrain and gain relevant skills, and ensure that existing inequalities are not exacerbated (PN 697).

In October 2023, the Government announced £118m to increase the UK's AI skills base.[40]

Workers' rights

The use of AI to manage work

Internationally, 'gig economy' [*] work such as taxi driving and food delivery is carried out through apps (such as Uber and Deliveroo[9,41,42]), which use AI systems to plan routes, determine how work is allocated, monitor worker performance, and determine

---

[*] The gig economy is defined by the Oxford English Dictionary as "a labour market characterised by a prevalence of short-term contracts and freelance work, as distinct from permanent, full-time employment".

pay.[9] Human intervention is currently still necessary to operate these AI systems, such as designing, maintaining, and troubleshooting algorithms (Table 1).

Some academics have raised concerns about wage discrimination[43,44] and power imbalances between workers and employers due to worker-generated data that employers can hold. They state that this can result in employers being able to undermine collective pay bargaining.[43,45,46]

There has also been an increase in AI used to manage office-based workers, particularly after the shift to remote and hybrid working during the Covid-19 pandemic. This includes surveillance of workers (PB 49), and use of AI in recruitment, such as sifting CVs.[47,48]

The Trades Union Congress[*] and APPG on the Future of Work, amongst others, have raised concerns about possible detrimental impacts on worker dignity and mental health due increased uses of AI for work allocation, monitoring, and disciplinary decisions.[9,47,49,50]

Concerns around labour practices of AI development

Training generative AI models may require datasets to be labelled,[51,52] which is often done manually (PB 57). Labelled data may be used for training generative AI models to identify harmful material and to not produce it.

Labelling data can often be outsourced by companies to workers in low- and middle-income countries. There have been some human rights concerns.[53–58] In 2023, media investigations by the Guardian, BBC and Time magazine amongst others into outsourced OpenAI data labellers in Kenya found workers were paid less than $2 an hour for labelling harmful content such as child sexual exploitation and violence.[58–60] This work is reportedly responsible for a detrimental impact on the workers' mental health.[58–60] In response, OpenAI said the outsourced workers could have opted out of this work "without penalisation".[58]

Use in public services

There are benefits and risks associated with the use of AI in public services, such as healthcare and education: the two largest public services by public spending.[61]

In healthcare, the use of AI could lead to better health outcomes by assisting with (PN 637):

- diagnosing diseases[4–7,62–69]

- devising personalised treatments[70,71]

- developing new drugs[72–74]

---

[*] The Trades Union Congress established a taskforce to write a draft AI and Employment Bill to be published in early 2024. The taskforce is advised by an expert committee consisting of technology industry group TechUK, the University of Oxford, the British Computer Society, trade unions, think tanks, and cross-party Parliamentarians.[50]

However, the expanding use of digital and AI technologies in healthcare may create barriers for digitally excluded communities, such as the elderly, in accessing healthcare.[75]

Whilst deployment is not yet widespread,[76] potential benefits from generative AI in education could include:

- bespoke educational courses that adapt to different learning styles

- reducing teachers' administrative work

- assisting teachers with feedback and marking[77]

However, AI may exacerbate existing inequalities due to differential access to AI technologies. There are concerns around students using AI in assessments, and there are risks around privacy breaches for teachers and students (see briefing on AI in education delivery and assessment).[78,79]

Algorithmic bias and discrimination

It has been well established that AI systems can have bias embedded into them,[80–87] which can manifest through various pathways,[88] including (PN 633):

1. Training datasets can be biased,[80,85–87] as they may consist of data generated and/or curated by humans with implicit or explicit bias.

2. Decisions made by humans in the design of algorithms, such as what attributes they want the algorithm to consider, may be implicitly or explicitly biased.

Widespread use of AI systems with unmitigated algorithmic bias could lead to discriminatory outcomes and exacerbated inequalities,[87,89,90] particularly in high-risk scenarios such as healthcare (PN 637).

For example, a 2019 study found that an algorithm used to allocate healthcare in US hospitals was less likely to refer Black people who were equally as sick as White people to healthcare programmes (PN 633).[91,92]

Such bias also has implications for human rights such as freedom from discrimination.

Responsibility, liability and transparency

The increasing use of automated decision-making raises implications for responsibility and liability.

It can be unclear to a person adversely affected by an automated decision that AI was used, what choices were made by developers, what went wrong, who is liable, and how to seek redress.[93–96] Depending on the type of incident, different parties, such as AI developers, deployers, or users (PB 57), could carry liability.[97–102]

Transparency on decisions allows individuals to know what has happened and exercise rights they may have.[103] In healthcare, medical ethicists have stated that responsible use of AI in diagnosis requires transparency, human oversight, and for

regulation to be clear about how liability is defined.[104–109] Many computer scientists and ethicists advocate for greater transparency when explaining how AI systems work ([PN 633]).[93,110–116]

For example, the Automated Vehicles Bill, which had its first reading in November 2023, provides drivers with immunity from prosecution relating to driving incidents if the self-driving vehicle is fully in control of itself, and places this liability on the company that created the vehicle.[117]

Misinformation and disinformation

Generative AI tools that generate inaccurate text[118–120] ([PB 57]) and realistic images, videos and other forms of mis- and disinformation* have become increasingly accessible.[121–124] Realistic images and videos generated for malicious purposes are commonly referred to as 'deepfakes'.

This accessibility has lowered the barrier for malicious actors to produce disinformation campaigns at scale,[123,125] although some academics[126] warn of exaggerated risks.

For example, in 2023, deepfake audio of Mayor of London Sadiq Khan saying Remembrance Day should be postponed was widely shared on social media.[127]

There are freedom of expression implications when regulating deepfakes, as generative AI can be used to create satirical content.[128–131] For example, artist Bruno Sartori produced a deepfake video criticising former Brazilian President Jair Bolsonaro's response to the Covid-19 pandemic.[131]

Elections, trust and engaging the public

Some experts say generative AI advances could increase public mistrust in online content, including election information,[132–134] and in institutions.[135–138] Others say concerns about the impact of fake images and news have been around for years.[136]

Some companies and news organisations are developing tools to let audiences know if content is AI generated, although technical challenges exist ([PB 57]).

Some experts say impacts of AI-based misinformation can be reduced through education in media literacy and fact-checking techniques.[136]

There are concerns around politicians using the atmosphere of distrust to discredit genuine evidence of their actions† by claiming that it is AI generated.[121,124,128,139,140]

---

* The UK Government defines disinformation as the "deliberate creation and spreading of false and/or manipulated information that is intended to deceive and mislead people, either for the purposes of causing harm, or for political, personal or financial gain". It defines misinformation as "the inadvertent spread of false information" ([PB 57]).

† Incidences of this have already been reported, such as an Indian politician who claimed a reportedly authentic audio clip of him was a deepfake.[140]

AI could also be used to strengthen democracy.[141] AI could be used to engage the public with politics and the electoral process. It could help voters understand manifestos and identify which candidates or political parties may best align with their priorities.[132,142]

Surveillance

There has been increasing use of both live and retrospective facial recognition[*] by private companies and police forces, as well as predictive policing that uses AI to predict hotspots for future crime.[143,144]

The Metropolitan Police said these tools save police officers' time, help identify criminals, and safeguard vulnerable people.[145]

Concerns have been raised by some academics,[146–154] parliamentarians,[155] human rights campaign groups,[155–158] and the Home Office Commissioner for Biometrics and Surveillance Cameras,[159] that live facial recognition, predictive policing,[160] and profiling[†] could restrict civil liberties and impact privacy.

The impacts could be real, in that authorities could limit freedom of expression and the right to protest, and/or perceived, in that individuals may impose restrictions on themselves due to an atmosphere of surveillance.[146,147]

Harassment, cybersecurity and scams

There have been numerous incidents where deepfake pornographic content of individuals, predominantly women,[161] has been shared online, leading to harassment, humiliation, and distress for individuals.[129,161–167] Sharing of non-consensual pornographic deepfakes has been criminalised by the Online Safety Act 2023.[168–170]

Generative AI can be used to create fake personas online[171], or impersonate real people.[172–176] This increases security risks, such as confidential information being unwittingly released to malicious actors, and convincing phishing and scam calls.[177]

---

[*] Live facial recognition refers to the use of facial recognition in real time via surveillance cameras. Retrospective facial recognition refers to the use of facial recognition after photographic or video evidence has been captured and/or taken from a different party. Facial recognition tools use AI (PB 57).

[†] Profiling is defined by the UK GDPR and ICO to be "any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements",[163] with profiling in the policing context referring to the aim of predicting an individual's propensity to crime.

Image ownership

AI can be used to recreate voices and images imitating living or deceased individuals.[178] This can have benefits in the arts, such as using AI for the consensual de-aging of performers.[179]

On the other hand, creative sector trade unions have raised concerns around companies being able to recreate the likeness of living or dead performers in perpetuity,[180,181] with implications for what fair remuneration for performers looks like.*

In the United States and many EU member states there exists a legal right to own your image.[182,183] There is no such right in the UK. However, privacy legislation, laws around misrepresentation, contract law, and other intellectual property rights could provide some protection for people who wish to control the use of their image.[183]

Some legal academics argue that the current body of law is not sufficient to protect people's image in the context of AI.[184]

The Government has committed to ratifying the Beijing Treaty on Audio-visual Performances†, which would give intellectual property rights to performers.[185]

Intellectual property

Generative AI tools are trained using datasets, which may or may not be open to the public (PB 57). Generative AI tools can output written, visual, aural, or audio-visual works that can mimic the style of specific human creators if their works are present in the datasets.[186] This raises implications for intellectual property rights, regardless of whether companies are transparent with datasets or not.[186–191]

In ongoing court cases, some authors and rightsholders have sued model developers in the UK and US alleging copyright infringement based on outputs that may imply that the datasets contain their work.[192,193]

There are also differing views on whether copyright should lie with users of AI tools, developers, those whose works appear in the dataset, or with nobody at all.[186,194,195]

Resource requirements

The Competition and Markets Authority,[196] US Federal Trade Commission,[197] and researchers,[198,199] have warned that the requirement of vast amounts of computing power for the largest AI models may restrict AI development to large companies and increase monopolisation in the technology sector.

---

* In 2023, implications for what fair remuneration looks like for performers came to the forefront as film and TV production in Hollywood ceased due to actors' and writers' strikes. Part of the disputes involved studio use of AI.[185]

† The Beijing Treaty was signed in 2013, but the UK Government were unable to ratify it independently while an EU member. The Intellectual Property Office has launched a consultation on the implementation of the Treaty.[189]

Concerns have been raised around the environmental impacts of intensive energy and water demands of AI infrastructure (PB 57).[200–202]

Existential risk to humanity

Some futurist philosophers and industry leaders have warned that AI may pose an existential risk to humanity if it were to supersede human thinking ability in every domain. Experts have varying views on the nature of future types of AI and what risks and opportunities it poses (PB 57).[203–207]

In 2023, some tech leaders, such as Elon Musk and Steve Wozniak, called for a six-month pause on the development of powerful AI to prioritise the mitigation of existential risks.[208,209]

A focus on existential risks has been heavily criticised by some academics,[210–215] the Ada Lovelace Institute,[216] and other industry experts[217–221] who say well-evidenced current risks should take precedence over speculative, long-term risks.[222,223] Some AI ethicists have raised concerns that focusing on existential risks diverts attention away from the decisions of tech leaders that are already affecting society.[214,224,225]

Others argue that existential risks should be taken seriously even if considered unlikely.[222]

Current regulatory environment

Leading AI companies operate and sell their products in multiple markets. The UK's regulatory environment must therefore also be seen in the context of other countries' regulatory regimes.

In the UK

While there is no current body of UK law specifically regulating AI, there are numerous laws that restrict how AI can be used in practice,[93,226] including (HoC Library briefing on AI and employment law):

- **data protection law**, such as the Data Protection Act 2018,[227] that affects data collection and processing for AI development, and is the remit of the Information Commissioner's Office

- **equalities, privacy and common law**, such as the Equality Act 2010[228] and the Human Rights Act 1998.[229] These laws affect the outcomes of AI systems and decisions which may have discrimination and human rights implications, and are the remit of the Equalities and Human Rights Commission. Privacy and common laws may limit the degree to which employers can substitute AI decision-making for their own judgement and places some restrictions on the use of surveillance tools to monitor workers[226]

- **intellectual property law**, such as the Copyright, Designs and Patents Act 1988,[230] which governs ownership and legal use of any intellectual property in outputs or in datasets, and is the remit of the Intellectual Property Office

The Digital Regulation Cooperation Forum (DRCF) was established in 2020 to foster collaboration between regulators regarding digital affairs.[231] As of 2023, the DRCF consists of Ofcom, the Information Commissioner's Office, the Competition and Markets Authority, and the Financial Conduct Authority.

In the 2023 white paper 'A pro-innovation approach to AI regulation',[21] the Government outlined a common set of cross-sectoral AI regulation principles for regulators to follow:

* safety, security, and robustness

* appropriate transparency and explainability

* fairness

* accountability and governance

* contestability and redress

In the EU and US

The European Union is currently finalising an AI Act.[232] As it stands, the Act is designed to work with existing EU legislation such as the General Data Protection Regulation (GDPR) and the Digital Services Act (DSA).

The Act defines different risk levels with corresponding levels of regulation. It bans certain high-risk applications, such as live facial recognition.

In the US, a 'Blueprint for an AI Bill of Rights' has been outlined.[233] Currently these are non-binding guidelines that aim to address discrimination, data privacy, and transparency. In October 2023, US President Joe Biden signed an Executive Order on AI mandating standards and disclosures for the largest AI companies, and measures to protect workers and disadvantaged groups.[234]

Potential future regulations

Human intervention in automated decision-making

Some civil society groups[235] and academics[236,237] have suggested a law to enshrine a right to human intervention in automated decision-making. There are two ways to implement this:

1. **All major decisions would be subject to human review.** Some legal scholars argue this is necessary so there is a designated human that can bear some degree of liability for the decision, while others argue this would reduce the efficiency gains of using automated decision-making.[235–240]

2. **Only contested decisions would be subject to human review.** Some legal scholars argue this is preferable for efficiency and sufficient to address

ethical concerns, while others argue that the lack of transparency may make it less likely that certain people, particularly those from digitally excluded communities, would be subject to a fair decision.[241]

Some legal academics have stated that mechanisms already exist in UK law (Section 49 of the Data Protection Act 2018)[227] to ensure human oversight in automated decision-making.[242]

However, other academics and civil society groups have raised concerns that in many systems, even when a human is involved in compliance with the law, they defer their judgement to AI.[239,242]

The Data Protection and Digital Information Bill,[243] carried into the 2023-2024 parliamentary session, which is intended by the Government to reduce burdens on businesses,[244] has been criticised by some civil society groups for "watering down protections"[245,246] present in current law against automated decision-making.

Ban on automated decision-making and ban on live facial recognition

Rather than regulating automated decision-making, some legal scholars propose banning it entirely and argue that it is an "illegitimate source of authority in liberal democracy", is not compatible with societal values such as equality and fairness, and that it can infer intimate knowledge about humans.* [247,248]

Civil liberty campaigners and some Parliamentarians have called for a ban on live facial recognition (see Surveillance)[155] akin to the ban proposed by the European Parliament in its draft AI Act.[232]

In 2020, the Equality and Human Rights Commission recommended the suspension of live facial recognition due to "discriminatory impacts".[249] The Metropolitan Police and Home Office have defended the use of live facial recognition, with the Home Office saying that it "has already enabled a large number of serious criminals to be caught".[250,251]

Open access to underlying AI code and related documentation

There has been considerable debate on whether companies developing AI models, particularly models used for automated decision-making, should make their models and documentation public and free to modify for transparency on how the models work. This may also promote competition by making AI developments accessible to small businesses.[252–257]

---

* An example of this was a shopping algorithm at US supermarket Target which could guess, based on insights from customers' shopping data, if a customer was pregnant. The algorithm would then send vouchers for pregnancy-related items, which in one case alerted a father to his daughter's teenage pregnancy.[251]

Companies such as OpenAI and Anthropic have argued against making AI code public, raising safety concerns around potential uses for malicious purposes.[258–261] Other companies (such as Meta and Mozilla) have voluntarily shared some of their underlying code.[262]

Algorithmic impact assessment and audits

Some learned societies and academics have said that a duty to carry out impact assessments of automated decisions (algorithmic impact assessments) could be placed on companies and public bodies.[263–268]

In existing regulation, the Data Protection Impact Assessment, mandated by Section 64 of the Data Protection Act 2018, places a duty on data controllers to assess data management processes that are "likely to result in a high risk to the rights and freedoms of individuals".[227,264]

Some stakeholders, such as the Institute for the Future of Work,[264] have stated algorithmic impact assessments could be modelled on the existing Data Protection Impact Assessment, and thus could involve:

- describing the AI systems and processes

- assessing the potential impact of AI processes on peoples' rights and freedoms

- measures to address those risks

- safeguards and mechanisms to ensure compliance with regulations

Algorithmic impact assessments exist in the Canadian public sector and require government agencies to complete a questionnaire before deploying automated decision-making.[269]

There is a voluntary Algorithmic Transparency Recording Standard that UK public bodies can use to disclose information about their use of AI.[270]

The Institute for the Future of Work has stated that audits of AI systems could ensure that these systems are compliant with their impact assessments and the law, and that the Digital Regulation Cooperation Forum could play a role in helping regulators cooperate on these audits.[264]

The Information Commissioners Office currently carries out consensual and compulsory data protection audits[*] of AI systems[271]. However, issues relating to AI are wider than data protection and fall under the remit of not just the Information Commissioners Office but multiple regulators (see Current regulatory environment).[264,272]

The Digital Regulation Cooperation Forum is currently exploring the future landscape of AI auditing.[272]

---

[*] Compulsory audits are carried out under Section 146 of the Data Protection Act, whereas consensual audits are carried out under Section 129 of the same act.

Opt-in/opt-out datasets

It has been proposed that an 'opt-in' or 'opt-out' model[*] could be used to give greater control to creatives over their works appearing in generative AI datasets.[194,273]

An opt-in model would prioritise creators. However, it may lead to restricted datasets due to low uptake that may hamper development. An opt-out model would automatically allow for more expansive datasets but increase the administrative burden on rightsholders and developers who must manually resolve opt-out requests.

The Government is working on a voluntary code of practice on copyright and AI.[274]

Regulatory capacity and funding

The Alan Turing Institute proposed a resource of expertise on AI that regulatory bodies could consult in order to respond to AI related matters that concern their individual remits.[275]

Experts and the Government have recommended to regulators the role of the Digital Regulation Cooperation Forum (see Current regulatory environment) could be expanded to be a central resource of expertise, and that more regulators (such as the Equalities and Human Rights Commission) could join to access this capacity.[21,252,275,276]

There are concerns amongst civil society[93,275,277,278] and Parliamentary committees,[252] that regulators are not currently equipped with the staffing, expertise or the funding to regulate AI and ensure current laws are enforced. Regulatory bodies[279] and the Government[21] have acknowledged these concerns.

---

[*] An opt-in model would require copyright holders to explicitly give consent for their intellectual property to be included in training datasets, whereas an opt-out model would mean that AI developers can use intellectual property by default unless rightsholders explicitly request to remove their work from the datasets.

References

1. Defence Science and Technology Laboratory (2020). AI, Data Science and Machine Learning: a Dstl biscuit book.

2. Busch, K. E. (2023). Social Media Algorithms: Content Recommendation, Moderation, and Congressional Considerations. *Congressional Research Service*.

3. Chen, B. X. *et al.* (2023). How Siri, Alexa and Google Assistant Lost the A.I. Race. *The New York Times*.

4. Dileep, G. *et al.* (2022). Artificial Intelligence in Breast Cancer Screening and Diagnosis. *Cureus*, Vol 14, e30318.

5. Marinovich, M. L. *et al.* (2023). Artificial intelligence (AI) for breast cancer screening: BreastScreen population-based cohort study of cancer detection. *eBioMedicine*, Vol 90, Elsevier.

6. Lång, K. *et al.* (2023). Artificial intelligence-supported screen reading versus standard double reading in the Mammography Screening with Artificial Intelligence trial (MASAI): a clinical safety analysis of a randomised, controlled, non-inferiority, single-blinded, screening accuracy study. *Lancet Oncol.*, Vol 24, 936–944. Elsevier.

7. Freeman, K. *et al.* (2021). Use of artificial intelligence for image analysis in breast cancer screening programmes: systematic review of test accuracy. *BMJ*, Vol 374, n1872. British Medical Journal Publishing Group.

8. Sample, I. (2017). AI watchdog needed to regulate automated decision-making, say experts. *The Guardian*.

9. Baiocco, S. (2022). The Algorithmic Management of work and its implications in different contexts. International Labour Organization.

10. Marsh, S. *et al.* (2020). Nearly half of councils in Great Britain use algorithms to help make claims decisions. *The Guardian*.

11. Department for Work and Pensions (2023). DWP annual report and accounts 2022 to 2023. *GOV.UK*.

12. British Computer Society (2022). DWP, machine algorithm and universal credit. *British Computer Society*.

13. Booth, R. (2023). AI use widened to assess universal credit applications and tackle fraud. *The Guardian*.

14. Seddon, P. (2023). Universal Credit: Warnings over AI use to risk-score benefit claims. *BBC News*.

15. Stacey, K. (2023). UK risks scandal over 'bias' in AI tools in use across public sector. *The Guardian*.

16. Li, S. *et al.* (2021). Planning and Decision-making for Connected Autonomous Vehicles at Road Intersections: A Review. *Chin. J. Mech. Eng.*, Vol 34, 133.

17. Liu, Q. *et al.* (2021). Decision-Making Technology for Autonomous Vehicles: Learning-Based Methods, Applications and Future Outlook. in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*.

18. Bommasani, R. *et al.* (2021). On the Opportunities and Risks of Foundation Models. Center for Research on Foundation Models (CRFM), Stanford Institute for

Human-Centered Artificial Intelligence (HAI).

19. Jones, E. Explainer: What is a foundation model? *Ada Lovelace Institute*.

20. Google Google Generative AI. *Google AI*.

21. Department for Science, Innovation and Technology (2023). AI regulation: a pro-innovation approach. *GOV.UK*.

22. Department for Science, Innovation and Technology (2023). Bristol set to host UK's most powerful supercomputer to turbocharge AI innovation. *GOV.UK*.

23. Foreign, Commonwealth & Development Office *et al.* (2023). AI Safety Summit 2023. *GOV.UK*.

24. Department for Science, Innovation and Technology *et al.* (2023). The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023.

25. Donelan, M. (2023). Introducing the AI Safety Institute. *GOV.UK*.

26. The Alan Turing Institute (online). Data science and AI glossary.

27. Cabinet Office *et al.* (2021). Ethics, Transparency and Accountability Framework for Automated Decision-Making. *GOV.UK*.

28. Department for Science, Innovation and Technology (2023). Artificial Intelligence sector study 2022. *GOV.UK*.

29. Green, A. (2023). Artificial intelligence and jobs: No signs of slowing labour demand (yet). OECD.

30. Hui, X. *et al.* (2023). The Short-Term Effects of Generative Artificial Intelligence on Employment: Evidence from an Online Labor Market. Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute.

31. Burn-Murdoch, J. (2023). Here's what we know about generative AI's impact on white-collar work. *Financial Times*.

32. Hayton, J. *et al.* (2023). Adoption of AI in UK firms - and the consequences for jobs. *Institute for the Future of Work*.

33. Briggs, J. *et al.* (2023). The Potentially Large Effects of Artificial Intelligence on Economic Growth. *Goldman Sachs*.

34. Department for Science, Innovation and Technology (2021). The potential impact of AI on UK employment and the demand for skills. *GOV.UK*.

35. Pawel Gmyrek, J. B. (2023). Generative AI and Jobs: A global analysis of potential effects on job quantity and quality. International Labour Organization.

36. Ellingrud, K. *et al.* (2023). Generative AI and the future of work in America. *McKinsey*.

37. Young, E. *et al.* (2023). Mind the gender gap: Inequalities in the emergent professions of artificial intelligence (AI) and data science. *New Technol. Work Employ.*, Vol 38, 391–414.

38. Hayton, J. *et al.* (2023). Briefing Paper: What drives UK firms to adopt AI and robotics, and what are the consequences for jobs? Zenodo.

39. Wall, J. *et al.* (2023). Making AI work for Britain. *techUK*.

40. Department for Science, Innovation and Technology, Britain to be made AI match-fit with £118 million skills package. *GOV.UK*.

41. Tuomi, A. *et al.* (2023). Riding Against the Algorithm: Algorithmic Management in On-Demand Food Delivery. in *Information and Communication Technologies in Tourism 2023*. (eds. Ferrer-Rosell, B. et al.) 28–39. Springer Nature Switzerland.

42. McDaid, E. *et al.* (2023). Algorithmic management and the politics of demand: Control and resistance at Uber. *Account. Organ. Soc.*, Vol 109, 101465.
43. Dubal, V. (2023). On Algorithmic Wage Discrimination. *Columbia Law Rev.*, Vol 123.
44. Cook, C. *et al.* (2021). The Gender Earnings Gap in the Gig Economy: Evidence from over a Million Rideshare Drivers. *Rev. Econ. Stud.*, Vol 88, 2210–2238.
45. Shapiro, A. (2020). Dynamic exploits: calculative asymmetries in the on-demand economy. *New Technol. Work Employ.*, Vol 35, 162–177.
46. Rosenblat, A. *et al.* (2016). Algorithmic Labor and Information Asymmetries: A Case Study of Uber's Drivers. *Int. J. Commun.*, Vol 10.
47. Trades Union Congress (2020). Technology managing people - The worker experience. *TUC*.
48. Nugent, S. E. *et al.* (2022). Recruitment AI Has a Disability Problem: Anticipating and Mitigating Unfair Automated Hiring Decisions. in *Towards Trustworthy Artificial Intelligent Systems*. (eds. Ferreira, M. I. A. et al.) 85–96. Springer International Publishing.
49. Milmo, D. (2021). Algorithmic tracking is 'damaging mental health' of UK workers. *The Guardian*.
50. APPG on Future of Work (2021). The New Frontier: Artificial intelligence at work. Institute for the future of work.
51. Amazon web services (online). What is data labeling?
52. IBM (online). What is Data Labeling?
53. Fraz, A. (2023). Hidden Workers powering AI. *National centre for AI*.
54. Hao, K. *et al.* (2022). How the AI industry profits from catastrophe. *MIT Technology Review*.
55. Aufiero, P. (2023). Pandora's Box: Generative AI Companies, ChatGPT, and Human Rights. Human Rights Watch.
56. Murgia, M. (2019). AI's new workforce: the data-labelling industry spreads globally. *Financial Times*.
57. Chandran, R. *et al.* (2023). FEATURE-AI boom is dream and nightmare for workers in Global South. *Reuters*.
58. Perrigo, B. (2023). Exclusive: The $2 Per Hour Workers Who Made ChatGPT Safer. *Time*.
59. Rowe, N. (2023). 'It's destroyed me completely': Kenyan moderators decry toll of training of AI models. *The Guardian*.
60. BBC News (2023). Kenyan AI worker traumatised from data labelling.
61. HM Treasury (2023). Public spending statistics: July 2023. *GOV.UK*.
62. Giorgio, J. *et al.* (2020). Modelling prognostic trajectories of cognitive decline due to Alzheimer's disease. *NeuroImage Clin.*, Vol 26, 102199.
63. Giorgio, J. *et al.* (2022). A robust and interpretable machine learning approach using multimodal biological data to predict future pathological tau accumulation. *Nat. Commun.*, Vol 13, 1887. Nature Publishing Group.
64. Barton, S. (2023). AI tool could speed up dementia diagnosis. *Sheffield University*.
65. Mirheidari, B. *et al.* (2021). Identifying Cognitive Impairment Using Sentence Representation Vectors. in *Interspeech 2021*. 2941–2945. ISCA.
66. Chaki, J. *et al.* (2022). Machine learning and artificial intelligence based Diabetes Mellitus detection

and self-management: A systematic review. *J. King Saud Univ. - Comput. Inf. Sci.*, Vol 34, 3204–3225.

67. Kaur, H. *et al.* (2020). Predictive modelling and analytics for diabetes using a machine learning approach. *Appl. Comput. Inform.*, Vol 18, 90–100. Emerald Publishing Limited.

68. Kaufman, J. M. *et al.* (2023). Acoustic Analysis and Prediction of Type 2 Diabetes Mellitus Using Smartphone-Recorded Voice Segments. *Mayo Clin. Proc. Digit. Health*, Vol 1, 534–544. Elsevier.

69. Kumar, Y. *et al.* (2023). Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda. *J. Ambient Intell. Humaniz. Comput.*, Vol 14, 8459–8486.

70. Schork, N. J. (2019). Artificial Intelligence and Personalized Medicine. in *Precision Medicine in Cancer Therapy*. (eds. Von Hoff, D. D. et al.) 265–283. Springer International Publishing.

71. Johnson, K. B. *et al.* (2021). Precision Medicine, AI, and the Future of Personalized Health Care. *Clin. Transl. Sci.*, Vol 14, 86–93.

72. Wong, F. *et al.* (2022). Benchmarking AlphaFold-enabled molecular docking predictions for antibiotic discovery. *Mol. Syst. Biol.*, Vol 18, e11081. John Wiley & Sons, Ltd.

73. Qureshi, R. *et al.* (2023). AI in drug discovery and its clinical relevance. *Heliyon*, Vol 9, e17575.

74. Paul, D. *et al.* (2021). Artificial intelligence in drug discovery and development. *Drug Discov. Today*, Vol 26, 80–93.

75. Studman, A. (2023). Access denied? Ada Lovelace Institute.

76. The Ada Lovelace Institute (online). Education and AI.

77. Cardona, M. A. *et al.* (2023). Artificial Intelligence and the Future of Teaching and Learning. *US Department of Education*.

78. The Bell Foundation (2023). Generative Artificial Intelligence in Education - Call for Evidence. *The Bell Foundation*.

79. O'Brien, N. *et al.* (2022). Addressing racial and ethnic inequities in data-driven health technologies. *53*. Institute of Global Health Innovation, Imperial College London.

80. West, S. M. (2019). Discriminating Systems: Gender, Race, and Power in AI - Report. AI Now Institute.

81. Department for Science, Innovation and Technology (2023). Frontier AI: capabilities and risks – discussion paper. *GOV.UK*.

82. Whittaker, M. *et al.* (2019). Disability, Bias, and AI - Report. AI Now Institute.

83. Ferrer, X. *et al.* (2021). Bias and Discrimination in AI: A Cross-Disciplinary Perspective. *IEEE Technol. Soc. Mag.*, Vol 40, 72–80.

84. Roselli, D. *et al.* (2019). Managing Bias in AI. in *Companion Proceedings of The 2019 World Wide Web Conference*. 539–544. Association for Computing Machinery.

85. Minderoo Centre for Technology and Democracy (2023). Definition drives design: Disability models and mechanisms of bias in AI technologies. *mctd.ac.uk*.

86. Buolamwini, J. *et al.* (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. in *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*. 77–91. PMLR.

87. Centre for Data Ethics and Innovation (2020). Review into

bias in algorithmic decision-making. *GOV.UK*.

88. International Organization for Standardization *et al.* (2021). Information technology — Artificial intelligence (AI) — Bias in AI systems and AI aided decision making.

89. Rotman, D. (2022). How to solve AI's inequality problem. *MIT Technology Review*.

90. Leslie, D. *et al.* (2021). Does "AI" stand for augmenting inequality in the era of covid-19 healthcare? *BMJ*, Vol 372, n304. British Medical Journal Publishing Group.

91. Obermeyer, Z. *et al.* (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, Vol 366, 447–453. American Association for the Advancement of Science.

92. Ledford, H. (2019). Millions of black people affected by racial bias in health-care algorithms. *Nature*, Vol 574, 608–609.

93. Davies, M. *et al.* (2023). Regulating AI in the UK. Ada Lovelace Institute.

94. Ogunleye, I. AI's Redress Problem: Recommendations to Improve Consumer Protection from Artificial Intelligence. *Center for Long-Term Cybersecurity*.

95. Public Law Project (2023). Government 'behind the curve' on AI risks.

96. Fanni, R. *et al.* (2023). Enhancing human agency through redress in Artificial Intelligence Systems. *AI Soc.*, Vol 38, 537–547.

97. Wendehorst, C. (2022). Liability for Artificial Intelligence: The Need to Address Both Safety Risks and Fundamental Rights Risks. in *The Cambridge Handbook of Responsible Artificial Intelligence: Interdisciplinary Perspectives*. (eds. Mueller, O. et al.) 187–209. Cambridge University Press.

98. Yeung, K. (2019). A Study of the Implications of Advanced Digital Technologies (Including AI Systems) for the Concept of Responsibility Within a Human Rights Framework. Council of Europe.

99. Bartneck, C. *et al.* (2021). Responsibility and Liability in the Case of AI Systems. in *An Introduction to Ethics in Robotics and AI*. (eds. Bartneck, C. et al.) 39–44. Springer International Publishing.

100. Maliha, G. *et al.* (2022). Who Is Liable When AI Kills? *Scientific American*.

101. Nix, M. *et al.* (2022). Understanding healthcare workers' confidence in AI. NHS AI Lab & Health Education England.

102. Buiten, M. *et al.* (2023). The law and economics of AI liability. *Comput. Law Secur. Rev.*, Vol 48, 105794.

103. Solomon, R. (2023). New data legislation will weaken our rights. *Public Law Project*.

104. World Health Organization (2021). Ethics and governance of artificial intelligence for health. *World Health Organization*.

105. Zhang, J. *et al.* (2023). Ethics and governance of trustworthy medical artificial intelligence. *BMC Med. Inform. Decis. Mak.*, Vol 23, 7.

106. Farhud, D. D. *et al.* (2021). Ethical Issues of Artificial Intelligence in Medicine and Healthcare. *Iran. J. Public Health*, Vol 50, i–v.

107. Gerke, S. *et al.* (2020). Ethical and legal challenges of artificial intelligence-driven healthcare. *Artif. Intell. Healthc.*, 295–336.

108. Murphy, K. *et al.* (2021). Artificial intelligence for good health: a scoping review of the ethics literature. *BMC Med. Ethics*, Vol 22, 14.

109. Neri, E. *et al.* (2020). Artificial intelligence: Who is responsible for the diagnosis? *Radiol. Med. (Torino)*, Vol 125, 517–521.

110. Larsson, S. *et al.* (2020). Transparency in artificial intelligence. *Internet Policy Rev.*, Vol 9.

111. Bagchi, S. (2023). What is a black box? A computer scientist explains what it means when the inner workings of AIs are hidden. *The Conversation*.

112. Savage, N. (2022). Breaking into the black box of artificial intelligence. *nature*.

113. Hassija, V. *et al.* (2023). Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence. *Cogn. Comput.*

114. Cassauwers, T. (2020). Opening the 'black box' of artificial intelligence | Research and Innovation. European Commission.

115. Blanc, R. M. and M. de (2023). Open Data and the AI Black Box. *Electronic Frontier Foundation*.

116. Wadden, J. J. (2022). Defining the undefinable: the black box problem in healthcare artificial intelligence. *J. Med. Ethics*, Vol 48, 764–768. Institute of Medical Ethics.

117. 10 Downing Street (2023). The King's Speech 2023: background briefing notes. *GOV.UK*.

118. Kumar, M. *et al.* (2023). Artificial Hallucinations by Google Bard: Think Before You Leap. *Cureus*, Vol 15, e43313.

119. Alkaissi, H. *et al.* (2023). Artificial Hallucinations in ChatGPT: Implications in Scientific Writing. *Cureus*, Vol 15, e35179.

120. Emsley, R. (2023). ChatGPT: these are not hallucinations – they're fabrications and falsifications. *Schizophrenia*, Vol 9, 1–2. Nature Publishing Group.

121. Kreps, S. *et al.* (2023). How AI Threatens Democracy. *J.*

122. Kreps, S. (2020). The role of technology in online misinformation. *Brookings*.

123. Funk, A. *et al.* (2023). The Repressive Power of Artificial Intelligence. *Freedom House*.

124. Chesney, R. *et al.* (2019). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. California Law Review.

125. Goldstein, J. A. *et al.* (2023). Generative Language Models and Automated Influence Operations: Emerging Threats and Potential Mitigations. *Stanford Internet Observatory*.

126. Simon, F. M. *et al.* (2023). Misinformation reloaded? Fears about the impact of generative AI on misinformation are overblown. Harvard Kennedy School Misinformation Review.

127. Sky News (2023). Deepfake audio of Sadiq Khan suggesting Remembrance weekend 'should be held next week instead' under police investigation. *Sky News*.

128. Kertysova, K. (2018). Artificial Intelligence and Disinformation: How AI Changes the Way Disinformation is Produced, Disseminated, and Can Be Countered. *Secur. Hum. Rights*, Vol 29, 55–81. Brill Nijhoff.

129. van der Sloot, B. *et al.* (2022). Deepfakes: regulatory challenges for the synthetic society. *Comput. Law Secur. Rev.*, Vol 46, 105716.

130. Barber, A. (2023). Freedom of expression meets deepfakes. *Synthese*, Vol 202, 40.

131. Gregory, S. (2021). JUST JOKING! Deepfakes, Satire, and the Politics of Synthetic Media. *MIT CoCreate*.

132. Adam, M. *et al.* (2023). Artificial intelligence, democracy and elections. *European Parliament*.

133. Quinn, B. *et al.* (2023). Time running out for UK electoral

Democr., Vol 34, 122–131. Johns Hopkins University Press.

system to keep up with AI, say regulators. *The Guardian*.

134. Panditharatne, M. *et al.* (2023). How AI Puts Elections at Risk — And the Needed Safeguards. *Brennan Center for Justice*.

135. Bantourakis, M. (2023). How can we build trustworthy media ecosystems in the age of AI and declining trust? *World Economic Forum*.

136. Kahn, G. (2023). Will AI-generated images create a new crisis for fact-checkers? Experts are not so sure. *Reuters Institute for the Study of Journalism*.

137. Mont'Alverne, C. *et al.* (2022). The trust gap: how and why news on digital platforms is viewed more sceptically versus news in general. *Reuters Institute for the Study of Journalism*.

138. Oxford Analytica (2023). Generative AI carries serious online risks.

139. Clementson, D. E. (2023). 6 ways AI can make political campaigns more deceptive than ever. *The Conversation*.

140. Kroetsch, J. (2023). Skepticism in Era of AI Deep Fakes Will Erode Defamation Claims. *Bloomberg Law*.

141. Krimmer, R. *et al.* (2022). Elections in digital times: a guide for electoral practitioners. *UNESCO*.

142. European Conferences of Electoral Management Bodies (2022). Concept paper 2022 - 'Artificial Intelligence and Electoral Integrity'. *European Conferences of Electoral Management Bodies*.

143. The Law Society (2019). Algorithm use in the criminal justice system report. *The Law Society*.

144. Babuta, A. *et al.* (2019). Data Analytics and Algorithmic Bias in Policing. *Royal United Services Institute*.

145. Metropolitan Police (online). Live Facial Recognition.

146. Fussey, P. *et al.* (2019). Independent Report on the London Metropolitan Police Service's Trial of Live Facial Recognition Technology. University of Essex Human Rights Centre.

147. Selinger, E. *et al.* (2019). The Inconsentability of Facial Surveillance. *Loyola Law Rev.*, Vol 66, 101–122.

148. Almeida, D. *et al.* (2022). The ethics of facial recognition technologies, surveillance, and accountability in an age of artificial intelligence: a comparative analysis of US, EU, and UK regulatory frameworks. *Ai Ethics*, Vol 2, 377–387.

149. Fontes, C. *et al.* (2021). Ethics of surveillance: harnessing the use of live facial recognition technologies in public spaces for law enforcement. Institute for Ethics in Artificial Intelligence, Technical University of Munich.

150. Mobilio, G. (2023). Your face is not new to me – Regulating the surveillance power of facial recognition technologies. *Internet Policy Rev.*, Vol 12.

151. Ovide, S. (2020). A Case for Banning Facial Recognition. *The New York Times*.

152. Murray, D. (2019). Live facial recognition: the impact on human rights and participatory democracy. *University of Essex*.

153. Minderoo Centre for Technology and Democracy (2022). A Sociotechnical Audit: Assessing Police use of Facial Recognition. *mctd.ac.uk*.

154. Ada Lovelace Institute (2019). Beyond face value: public attitudes to facial recognition technology.

155. Big Brother Watch Team (2023). 65 parliamentarians call for "immediate stop" to live facial

recognition surveillance. *Big Brother Watch*.

156. Liberty (2021). Human Rights coalition calls for immediate ban on facial recognition. *Liberty*.

157. Akram, S. (2023). UK Facial Recognition – No Consent, No Oversight. *Open Rights Group*.

158. Collings, P. *et al.* (2022). Ban Government Use of Face Recognition In the UK. *Electronic Frontier Foundation*.

159. Boffey, D. *et al.* (2023). Britain is 'omni-surveillance' society, watchdog warns. *The Guardian*.

160. Couchman, H. (2019). Report: Policing by machine. *Liberty*.

161. Laffier, J. *et al.* (2023). Deepfakes and Harm to Women. *J. Digit. Life Learn.*, Vol 3, 1–21.

162. Flynn, A. *et al.* (2022). Deepfakes and Digitally Altered Imagery Abuse: A Cross-Country Exploration of an Emerging form of Image-Based Sexual Abuse. *Br. J. Criminol.*, Vol 62, 1341–1358.

163. Flynn, A. *et al.* (2021). Disrupting and Preventing Deepfake Abuse: Exploring Criminal Law Responses to AI-Facilitated Abuse. in *The Palgrave Handbook of Gendered Violence and Technology*. (eds. Powell, A. et al.) 583–603. Springer International Publishing.

164. Graber-Mitchell, N. (2021). Artificial Illusions: Deepfakes as Speech. *Intersect*, Vol 14.

165. Hao, K. (2021). Deepfake porn is ruining women's lives. Now the law may finally ban it. *MIT Technology Review*.

166. Minderoo Centre for Technology and Democracy (2022). Tackling AI-enabled intimate image abuse. *mctd.ac.uk*.

167. Ajder, H. *et al.* (2019). The State of Deepfakes: Landscape, Threats, and Impact. Deeptrace.

168. Online Safety Act 2023. King's Printer of Acts of Parliament.

169. Rahman-Jones, I. *et al.* (2023). Online Safety Bill: divisive internet rules become law. *BBC News*.

170. Department for Science, Innovation and Technology (2022). A guide to the Online Safety Bill. *GOV.UK*.

171. Cross, C. (2022). Using artificial intelligence (AI) and deepfakes to deceive victims: the need to rethink current romance fraud prevention messaging. *Crime Prev. Community Saf.*, Vol 24, 30–41.

172. Wise, J. (2023). Imagine your child calling for money. Except it's not them – it's an AI scam. *The Guardian*.

173. Hughes, A. (2023). AI: Why the next call from your family could be a deepfake scammer. *BBC Science Focus Magazine*.

174. Buckley, O. (2023). AI scam calls imitating familiar voices are a growing problem – here's how they work. *The Conversation*.

175. David, D. (2021). Analyzing The Rise Of Deepfake Voice Technology. *Forbes*.

176. Green, J. *et al.* (2023). Experts warn of rise in scammers using AI to mimic voices of loved ones in distress. *ABC News*.

177. Brundage, M. *et al.* (2018). The Malicious Use of Artificial Intelligence. *Malicious AI Report*.

178. Leffer, L. (2023). Can AI Replace Actors? Here's How Digital Double Tech Works. *Scientific American*.

179. Smith, D. (2023). 'We're going through a big revolution': how AI is de-ageing stars on screen. *The Guardian*.

180. Equity (online). AI Vision Statement.

181. Maddaus, G. (2023). SAG-AFTRA Strike: AI Fears Mount for Background Actors. *Variety*.

182. Barbas, S. (2015). Laws of Image: Privacy and Publicity in America. Stanford University Press.

183. HM Revenue & Customs (2023). HMRC internal manual Capital Gains Manual.

184. Bosher, H. (2023). Forced Faming: How the Law Can Protect Against Non-Consensual Recording and Distributing of a Person's Voice or Image. *Commun. Law*, Vol 28, 119–125.

185. Intellectual Property Office (2023). Government consults on implementation of Beijing Treaty on Audiovisual Performances. *GOV.UK*.

186. Epstein, Z. *et al.* (2023). Art and the science of generative AI. *Science*, Vol 380, 1110–1111. American Association for the Advancement of Science.

187. Armstrong, H. *et al.* (2023). Generative AI and intellectual property rights—the UK government's position. *Reynolds Porter Chamberlain*.

188. Appel, G. *et al.* (2023). Generative AI Has an Intellectual Property Problem. *Harvard Business Review*.

189. Smits, J. *et al.* (2022). Generative AI and Intellectual Property Rights. in *Law and Artificial Intelligence: Regulating AI and Applying AI in Legal Practice*. (eds. Custers, B. et al.) 323–344. T.M.C. Asser Press.

190. Chesterman, S. (2023). Good Models Borrow, Great Models Steal: Intellectual Property Rights and Generative AI. National University of Singapore.

191. Baker, C. (2023). Intellectual property and generative AI. *Deloitte*.

192. Italie, H. (2023). 'Game of Thrones' creator and other authors sue ChatGPT-maker OpenAI for copyright infringement. *AP News*.

193. Getty Images (2023). Getty Images Statement. *Getty Images*.

194. Bosher, H. (2023). Policy Brief: Copyright, generative AI and data mining. *Brunel University London*.

195. Brittain, B. (2023). AI-generated art cannot receive copyrights, US court says. *Reuters*.

196. Competition and Markets Authority (2023). AI Foundation Models: Initial report. *GOV.UK*.

197. Crawford, K. (2023). FTC's Lina Khan warns Big Tech over AI. *Stanford Institute for Economic Policy Research (SIEPR)*.

198. West, S. M. (2023). Competition authorities need to move fast and break up AI. *Financial Times*.

199. Bengio, Y. *et al.* (2018). Countering the monopolization of research. *UNESCO*.

200. Ligozat, A.-L. *et al.* (2022). Unraveling the Hidden Environmental Impacts of AI Solutions for Environment Life Cycle Assessment of AI Solutions. *Sustainability*, Vol 14, 5172. Multidisciplinary Digital Publishing Institute.

201. Stokel-Walker, C. (2023). TechScape: Turns out there's another problem with AI – its environmental toll. *The Guardian*.

202. Hsu, J. (2023). Shifting where data is processed for AI can reduce environmental harm. New Scientist.

203. Müller, V. C. *et al.* (2016). Future Progress in Artificial Intelligence: A Survey of Expert Opinion. in *Fundamental Issues of Artificial Intelligence*. (ed. Müller, V. C.) 555–572. Springer International Publishing.

204. Bubeck, S. *et al.* (2023). Sparks of Artificial General Intelligence: Early experiments with GPT-4. arXiv.

205. Metz, C. (2023). Microsoft Says New AI Shows Signs of Human Reasoning. *The New York Times*.

206. Nathan, A. *et al.* (2023). Generative AI: hype, or truly transformative? *Goldman Sachs*.

207. The Academy of Medical Sciences *et al.* (2023). A joint statement on the AI Safety Summit.

208. Future of Life Institute (2023). Pause Giant AI Experiments: An Open Letter. *Future of Life Institute*.

209. Center for AI Safety (2023). Statement on AI Risk. *Center for AI Safety*.

210. Eisikovits, N. (2023). AI is an existential threat – just not the way you think. *The Conversation*.

211. Nature Editorial (2023). Stop talking about tomorrow's AI doomsday when AI poses risks today. *Nature*, Vol 618, 885–886.

212. Coldewey, D. (2023). Ethicists fire back at 'AI Pause' letter they say 'ignores the actual harms'. *TechCrunch*.

213. Gebru, T. *et al.* (2023). Statement from the listed authors of Stochastic Parrots on the "AI pause" letter. *DAIR Institute*.

214. Science Media Centre (2023). Expert reaction to a statement on the existential threat of AI published on the Centre for AI Safety website. *Science Media Centre*.

215. Richards, B. *et al.* (2023). The Illusion Of AI's Existential Risk. the Berggruen Institute.

216. Davies, M. *et al.* (2023). Seizing the 'AI moment': making a success of the AI Safety Summit. Ada Lovelace Institute.

217. Heikkilä, M. (2023). Meta's AI leaders want you to know fears over AI existential risk are "ridiculous". *MIT Technology Review*.

218. Hammond, G. (2023). Aidan Gomez: AI threat to human existence is 'absurd' distraction from real risks. *Financial Times*.

219. Tucker, I. (2023). Signal's Meredith Whittaker: 'These are the people who could actually pause AI if they wanted to'. *The Observer*.

220. Heaven, W. D. (2023). How existential risk became the biggest meme in AI. *MIT Technology Review*.

221. Arcas, B. A. y (2023). Fears about AI's existential risk are overdone, says a group of experts. *The Economist*.

222. Sætra, H. S. *et al.* (2023). Resolving the battle of short- vs. long-term AI risks. *AI Ethics*.

223. Vynck, G. D. (2023). The debate over whether AI will destroy us is dividing Silicon Valley. *Washington Post*.

224. Aitken, M. (2023). The real reason claims about the existential risk of AI are scary. *New Scientist*.

225. Cremer, C. Z. *et al.* (2021). Democratising Risk: In Search of a Methodology to Study Existential Risk.

226. Brione, P. (2023). Artificial intelligence and employment law.

227. Data Protection Act 2018. King's Printer of Acts of Parliament.

228. Equality Act 2010. Statute Law Database.

229. Human Rights Act 1998. Statute Law Database.

230. Copyright, Designs and Patents Act 1988. Statute Law Database.

231. Competition and Markets Authority *et al.* (2023). The Digital Regulation Cooperation Forum. *GOV.UK*.

232. European Parliament (2023). EU AI Act: first regulation on artificial intelligence. *European Parliament*.

233. Blueprint for an AI Bill of Rights | OSTP. *The White House*.

234. The White House (2023). FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence. *The White House*.

235. Andrews, E. (2023). Liberty's Written Submission to a pro-innovation approach to AI regulation consultation. *Liberty*.

236. Dror-Shpoliansky, D. *et al.* (2020). It's the End of the (Offline) World as We Know It: from Human Rights to Digital Human Rights – a Proposed Typology. Hebrew University of Jerusalem.

237. Almada, M. (2019). Human intervention in automated decision-making: Toward the construction of contestable systems. in *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Law*. 2–11. Association for Computing Machinery.

238. Binns, R. (2022). Human Judgment in algorithmic loops: Individual justice and automated decision-making. *Regul. Gov.*, Vol 16, 197–211.

239. Lazcoz, G. *et al.* (2023). Humans in the GDPR and AIA governance of automated and algorithmic systems. Essential pre-requisites against abdicating responsibilities. *Comput. Law Secur. Rev.*, Vol 50, 105833.

240. Enarsson, T. *et al.* (2022). Approaching the human in the loop – legal perspectives on hybrid human/algorithmic decision-making in three contexts. *Inf. Commun. Technol. Law*, Vol 31, 123–153. Routledge.

241. Cohen, I. G. *et al.* (2023). How AI can learn from the law: putting humans in the loop only on appeal. *Npj Digit. Med.*, Vol 6, 1–4. Nature Publishing Group.

242. Abrusci, E. *et al.* (2023). The questionable necessity of a new human right against being subject to automated decision-making. *Int. J. Law Inf. Technol.*, Vol 31, 114–143.

243. House of Commons (2023). Data Protection and Digital Information Bill. *UK Parliament*.

244. Department for Digital, Culture, Media & Sport (2022). Data: a new direction - government response to consultation. *GOV.UK*.

245. Trades Union Congress (2023). Briefing: Data Protection and Digital Information Bill (V2). *TUC*.

246. Yuill, B. (2023). How the new Data Bill waters down protections. *Public Law Project*.

247. Hill, K. (2012). How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did. *Forbes*.

248. Waldman, A. E. (2019). Power, Process, and Automated Decision-Making. *Fordham Law Rev.*, Vol 88, 613–632.

249. Equality and Human Rights Commission (2020). Civil and political rights in Great Britain: submission to the UN. *EHRC*.

250. Rahman-Jones, I. *et al.* (2023). AI facial recognition: Campaigners and MPs call for ban. *BBC News*.

251. Science, Innovation and Technology Committee (2023). Science, Innovation and Technology Committee Oral evidence: Governance of artificial intelligence (AI), HC 945. *House of Commons*.

252. Science, Innovation and Technology Committee The governance of artificial intelligence: interim report. *House of Commons*.

253. Shrestha, Y. R. *et al.* (2023). Building open-source AI. *Nat. Comput. Sci.*, 1–4. Nature Publishing Group.

254. Manancourt, V. (2023). British deputy PM throws backing behind open source AI. *POLITICO*.

255. Bostrom, N. (2018). Strategic Implications of Openness in AI Development. in *Artificial Intelligence Safety and Security*. Chapman and Hall/CRC.

256. Ricks, B. *et al.* (2020). Creating Trustworthy AI. *Mozilla Foundation*.

257. Ding, J. (2022). What defines the 'open' in 'open AI'? *The Alan Turing Institute*.

258. Vincent, J. (2023). OpenAI co-founder on company's past approach to openly sharing research: "We were wrong". *The Verge*.

259. Tett, G. (2023). The perils of open-source AI. *Financial Times*.

260. Financial Times (2023). AI: open-source models imperil profits of big tech's contenders. *Financial Times*.

261. Broughel, J. (2023). How Regulating AI Could Empower Bad Actors. *Forbes*.

262. Metz, C. *et al.* (2023). In Battle Over A.I., Meta Decides to Give Away Its Crown Jewels. *The New York Times*.

263. Ada Lovelace Institute (2020). Examining the Black Box. Ada Lovelace Institute.

264. Institute for the Future of Work (2021). Algorithmic Impact Assessments. *Institute for the Future of Work*.

265. Groves, L. (2022). Algorithmic impact assessment: a case study in healthcare. *Ada Lovelace Institute*.

266. Kaminski, M. E. *et al.* (2021). Algorithmic impact assessments under the GDPR: producing multi-layered explanations. *Int. Data Priv. Law*, Vol 11, 125–144.

267. Selbst, A. D. (2021). An Institutional View Of Algorithmic Impact Assessments. *Harv. J. Law Technol.*, Vol 35, 117–191.

268. Reisman, D. *et al.* (2018). Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability. AI Now Institute.

269. Government of Canada (2023). Algorithmic Impact Assessment Tool.

270. Central Digital & Data Office *et al.* (2023). Algorithmic Transparency Recording Standard Hub. *GOV.UK*.

271. Information Commissioner's Office (online). A Guide to ICO Audit Artificial Intelligence (AI) Audits. Information Commissioner's Office.

272. Digital Regulation Cooperation Forum (2022). Auditing algorithms: the existing landscape, role of regulators and future outlook. *GOV.UK*.

273. Intellectual Property Office (2022). Artificial Intelligence and Intellectual Property: copyright and patents: Government response to consultation. *GOV.UK*.

274. Intellectual Property Office (2023). The government's code of practice on copyright and AI. *GOV.UK*.

275. Aitken, M. *et al.* (2022). Common Regulatory Capacity for AI. The Alan Turing Institute.

276. Jung, C. *et al.* (2023). Artificial intelligence for public value creation. *Inst. Public Policy Res.*,

277. Roberts, H. *et al.* (2023). Artificial intelligence regulation in the United Kingdom: a path to good governance and global leadership? *Internet Policy Rev.*, Vol 12.

278. Minderoo Centre for Technology and Democracy (2023). Policy Brief: Generative AI. *mctd.ac.uk*.

279. Equality and Human Rights Commission (2023). AI safeguards 'inadequate', watchdog warns. *EHRC*.

Contributors

Members of the POST Board*

Dr Elena Abrusci, Brunel University London*

Dr Mhairi Aitken, Alan Turing Institute*

Emmanuelle Andrews, Liberty

Dr Hayleigh Bosher, Brunel University London*

Matt Davies, Ada Lovelace Institute

Maximilian Gahntz, Mozilla Foundation

Conor Griffin, Google DeepMind

Professor Oliver Hauser, University of Exeter

Harry Law, Google*

Mia Leslie, Public Law Project*

Mavis Machirori, Ada Lovelace Institute

Professor Gina Neff, University of Cambridge

Sam Nutt, London Office of Technology and Innovation

Lucy Purdon, Mozilla Foundation*

Adam Smith, British Computer Society*

Amy Smith, Queen Mary University of London

Anna Studman, Ada Lovelace Institute

Mary Towers, Trades Union Congress

Professor Shannon Vallor, University of Edinburgh*

National Centre for AI for Tertiary Education, Jisc

*denotes people and reviewers who acted as external reviewers of the briefing

Image Credit: Image by Romana from Adobe Stock

✉@ post@parliament.uk

🌐 parliament.uk/post

🐦 @POST_UK