AI AND AUTOMATED DECISION MAKING
TO REGULATE OR DEREGULATE?

Alexander Dittel of Wedlake Bell LLP discusses the current and future regulation of artificial intelligence and automated decision-making technologies.

Over the last ten years, algorithmic decisions have slowly found their way into daily life. Connected devices, digitalisation and digital transformation, as well as advancements in graphics processing units, have enabled technologies to become personal assistants, public servants and much more. However, not everyone is ready to trade their personal privacy for the benefits of artificial intelligence (AI). AI cannot function without massive amounts of data, and AI deployment and outputs can give rise to significant human rights concerns and ethical issues.

As Professor Stephen Hawking said, "the rise of powerful artificial intelligence will be either the best, or the worst, thing ever to happen to humanity" (*www.cam.ac.uk/research/news/the-best-or-worst-thing-to-happen-to-humanity-stephen-hawking-launches-centre-for-the-future-of*). "It could also be the last, unless we learn how to avoid the risks." This leads to the question of whether current regulation is sufficient to protect individuals' rights, including the right to privacy. Conversely, the global race for AI supremacy, with its huge economic potential, is likely to be won by countries with lighter regulation.

This article discusses the current regulation of AI and automated decision making (ADM), the key concerns and risks that arise with the use of AI and ADM systems, and the future of AI and ADM regulation, including the government's recent proposals on data protection reform.

REGULATING AI

In 2019, KPMG found that only 17% of companies interviewed reported the use of AI or machine learning at scale (*https://advisory.kpmg.us/articles/2019/ai-transforming-enterprise.html*). However, in February 2021, KPMG predicted that 2021 would herald AI's unprecedented leap to the mainstream, largely as a result of the COVID-19 pandemic (*https://home.kpmg/xx/en/home/insights/2021/02/trends-in-artificial-intelligence.html*).

Recently, Google's AlphaFold 2 contributed to resolving the protein folding problem by modelling the three-dimensional structure of amino acids within hours, as opposed to the months it would take with crystallography. Tesla's full self-driving beta technology is coming ever closer to tackling the complexities of the open road. Although today's AI excels in narrowly defined uses rather than complex tasks, the extraordinary achievements of AI continue to impress and its applications are becoming ever wider (*see box "Artificial intelligence systems"*).

The recent meteoric rise in the use of AI has brought to the fore practical and ethical issues around its use. For example, the

Irish Council for Civil Liberties has brought an action against adtech industry trade body, IAB TechLab, alleging that, through real-time bidding, companies are building secret dossiers of each online user's intimate characteristics and broadcasting these to thousands of companies, hundreds of billions of times, every day (*www.iccl.ie/rtb-june-2021*). These huge datasets are then processed by complex algorithms to instantly serve "relatable" advertising content to audiences of the highest bidder.

On 24 November 2021, UNESCO (the United Nations Educational, Scientific and Cultural Organization) adopted a recommendation on the ethics of AI, which calls for an inclusive and collaborative approach (*see box "UNESCO recommendation"*). AI that is heavily influenced by local cultural norms may not serve well elsewhere in the world. Local cultural norms can be embraced as long as they are consistent with core values and principles, which must prevail over the human behaviour that AI is hoped to imitate. Many would argue that the importance of the UK's strong voice in building global ethical standards for AI cannot be underestimated.

In light of these issues, the question arises whether Article 22 of the General Data Protection Regulation (*679/2016/EU*) (GDPR) (Article 22) or the concept of legitimate interests is sufficient to protect people when AI is used to predict behaviour to encourage consumer spending, tailor content to keep people online, compare candidates' job interviews, decide eligibility for private or public services, or evaluate work performance within a fraction of a second (*see News brief "Legitimate interests under the GDPR: flexibility but at a cost", www.practicallaw.com/w-014-5328*).

Privacy advocates around the world are calling for further regulation and the EU is finalising its Artificial Intelligence Act, which will establish a harmonised legal framework for the regulation of AI (*see News brief "Artificial intelligence: the dawn of a new legal era", www.practicallaw.com/w-031-0858*). The EU has gone even further by proposing a directive on improving working conditions in digital labour platforms on 9 December 2021 (the draft Digital Labour Platforms Directive) (*https://ec.europa.eu/commission/presscorner/detail/en/ip_21_6605*). However, judging by its consultation on proposals to reform UK data protection, which was published on 10 September 2021, the government is leaning towards deregulation (*see "UK data reform" below*). After all, AI promises to unlock an estimated $15.7 trillion in new economic value to global GDP by 2030, according to a report by PwC (*www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html*).

The global race for AI supremacy favours the likes of China and the US, which have little regulation of AI. The government is poised not to let the GDPR cost the UK its position in that race and has proposed removing Article 22 from the retained EU law version of the GDPR (UK GDPR). Rewriting the legal rules of AI, with the promise of AI rewriting the rules of industry, might help to make the UK an AI superpower.

KEY CONCERNS

AI learns from labelled datasets and absorbs all prejudices, misconceptions and failures into its algorithmic decision making, which could have a profound effect on human rights. This risk is proliferated by encouraging vast data collection for instantaneous automated processing and decision making on a large scale. Due to the high risk, AI outputs must be benchmarked against the noblest examples of humanity and abstract ethical principles.

As AI is deployed to make increasingly significant decisions about human life, skilled humans are needed to challenge those decisions, which are often probabilistic and may be biased. The use of human intervention is an essential safeguard. However, AI systems are often used to replace skilled workers and save costs. Without transparency, meaningful human intervention and data protection rights, AI could significantly interfere with human rights.

In the absence of global standardisation, AI architects are at liberty to decide what data AI should process and what outcomes are appropriate. For example, a collection of nine Chinese state agencies have mandated that AI should carry forward socialist core values and uphold people's "correct" political direction and public opinion orientation (*Chinese Guiding Opinions on Strengthening Overall Governance of Internet Information Service Algorithms, 17 September 2021, translated by the DigiChina Project, Stanford University, https://digichina.stanford.edu/work/translation-guiding-opinions-on-strengthening-overall-governance-of-internet-information-service-algorithms/*). There is a risk that ideological goals such as these could quickly turn humanity into a servant of AI, having a significant impact on human rights, diversity and pluralism, cultural expression and equality.

Equally concerning is the Western world's use of AI to influence consumer choices. In addition to privacy and data protection concerns, AI technology can be used to promote discrimination, prejudice and harmful ideologies (*see feature article "Algorithms, apps and AI: the next frontier*

Artificial intelligence systems

Artificial intelligence (AI) systems are information-processing technologies that integrate models and algorithms to produce a capacity to learn and to perform cognitive tasks, leading to outcomes such as prediction and decision making without human involvement.

Today's so-called "narrow AI" has the ability to learn specific tasks, such as translating language, recognising speech or spotting anomalies on x-rays, without being explicitly programmed for those tasks. Neural networks that mimic brain neurons allow interconnected layers of algorithms to communicate and assign scores to data that passes between these layers. A task is learned when the output becomes accurate or close to accurate. Massive amounts of labelled datasets, as well as extraordinary computing power, are essential for training AI.

The next level up is artificial general intelligence or AGI, which refers to the hypothetical ability of software to find solutions and acquire new skills when faced with unfamiliar tasks. It will start to resemble human self-consciousness and emotions, with the ability to think abstractly and recall its thoughts and memories to come up with informed decisions, strategies and even creative ideas. This flexible form of AI is not expected to exist until at least 2050.

in discrimination law", www.practicallaw.com/w-013-8054). More importantly, today's lucrative automated journalism and the algorithmic provision of news, and the moderation and curation of content on social media and search engines, gives rise to:

- Misinformation; that is, untrue or misleading information that is presented as fact, regardless of the intent to mislead, such as believing a false rumour on social media and spreading it to others.
- Disinformation; that is, the deliberate dissemination of false information in order to deceive, mislead or manipulate people.
- Hate speech; that is, forms of expression that promote hatred, violence or discrimination against a person or group of people.
- The emergence of new forms of societal narratives that affect the population's critical thinking and information literacy skills.
- The creation of echo chambers that increase social and political polarisation and extremism.

In the long term, AI systems could challenge humans' special sense of existence, self-understanding, interaction, autonomy, worth and dignity. AI could eradicate the valuable lessons of making one's own mistakes and lead human lives on a meticulously pre-arranged path. Some suggest that the way forward will be creating a synergy between the human brain and AI in order to achieve harmony.

If laws exist because human nature cannot be trusted, then surely AI, which copies human behaviour, cannot be trusted either. In 2018, SpaceX and Tesla CEO Elon Musk warned that AI "doesn't have to be evil to destroy humanity. If AI has a goal and humanity just happens to come in the way, it will destroy humanity as a matter of course without even thinking about it" (*www.cnbc.com/2018/04/06/elon-musk-warns-ai-could-create-immortal-dictator-in-documentary.html*).

ARTICLE 22

Article 22 regulates ADM, which is often integral to AI systems. AI is used to assist a human decision maker or to make decisions automatically. However, Article 22 applies only if the ADM has a significant effect on an individual and there is no human intervention in the decision-making process. It provides that data subjects have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning them or similarly significantly affects them. If triggered, Article 22 imposes additional requirements on transparency, lawful basis and safeguards in relation to ADM, such as the right to obtain human intervention and contest the decision.

Uber and Ola

In March 2021, Uber and Ola drivers in the Netherlands partially succeeded in establishing that they are employees rather than independent contractors by arguing that they are subject to ADM, which is inconsistent with the status of an independent contractor (*Uber drivers v Uber, 11 March 2021, Amsterdam*

UNESCO recommendation

The UNESCO recommendation on the ethics of artificial intelligence (AI) is based on a holistic, comprehensive, multicultural and evolving framework of interdependent values, principles and actions that can guide societies in dealing responsibly with the known and unknown effects of AI technologies on humanity, offering a basis to accept or reject AI technologies (*https://en.unesco.org/artificial-intelligence/ethics*). AI ethics are rooted in the ethics of science and technology, and is a dynamic basis for evaluating AI technologies in order to realise the advantages of technology while promoting human dignity and wellbeing, and reducing human rights risks associated with its use.

The recommendation suggests that every AI system should promote certain core values that apply worldwide, such as human dignity, rights and freedoms, the protection of the environment and ecosystem, diversity and inclusiveness, and peaceful and interconnected societies. These values are underpinned by the ten principles of:

- Proportionality and doing no harm.
- Safety and security.
- Fairness and non-discrimination.
- Sustainability.
- Right to privacy and data protection.
- Human oversight and determination.
- Transparency and explainability.
- Responsibility and accountability.
- Awareness and literacy.
- Multi-stakeholder and adaptive governance and collaboration.

Given China's global dominance in AI, its adoption of the recommendation is very positive. However, the recommendation is not binding. According to a recent report, China is building the world's most sophisticated surveillance technology network, which is intended to track journalists and foreign students, along with other "suspicious" individuals (*www.reuters.com/technology/exclusive-chinese-province-targets-journalists-foreign-students-with-planned-new-2021-11-29/*). While the recommendation embraces the diversity of cultural systems, which are essential for building sustainable societies around the world, it also discourages the use of AI for social scoring and mass surveillance purposes.

District Court, C/13/687315; Uber drivers v Uber, 11 March 2021, Amsterdam District Court, C/13/692003; Ola drivers v Ola Cabs, 11 March 2021, Amsterdam District Court, C/13/689705).

The Amsterdam District Court held that that ADM relating to matching passengers with drivers, assigning trips, maintaining drivers' earning profile and detecting irregularities did not legally affect or similarly significantly affect them under Article 22. However, it held that Ola's automated system of penalties and deductions in respect of invalid rides did have a significant effect because it affected the drivers' contractual rights.

Uber successfully avoided the application of Article 22 by arguing that its algorithm merely provides information about suspected fraud and the decision to exclude a driver is made by the human operational risk team, rather than ADM. However, according to the European Data Protection Board's (EDPB) guidance on automated individual decision making and profiling for the purposes of the GDPR, human intervention must be meaningful (*https://ec.europa.eu/newsroom/article29/items/612053/en*). Merely rubberstamping an automatically made decision will not suffice to avoid Article 22.

This kind of circumvention will likely not be possible under the draft Digital Labour Platforms Directive, which will impose transparency, human intervention and data minimisation obligations in relation to any automated monitoring and ADM systems that could significantly affect working conditions. This will include ADM about access to work assignments, earnings, occupational safety and health, working time, promotion, contractual status, and restriction, suspension or termination of users' accounts.

The court clarified that, if Article 22 is triggered, the controller must provide meaningful information to individuals about the logic involved in the ADM. This will include the main assessment criteria and their role in the automated decision in order to enable individuals to understand the basis of ADM and to verify its correctness and lawfulness.

EDP Energia

As reported by the EDPB in May 2021, the Spanish data protection agency (DPA) imposed a penalty of €1.5 million on utilities company EDP Energia SAU for relying on a consent for ADM that was held to be invalid (*https://edpb.europa.eu/news/national-news/2021/spanish-dpa-imposes-fine-1500000-euros-epd-energia-sau-two-infractions-gdpr_en*). The consent wording and privacy notice bundled profiling for advertising and ADM about the customer's eligibility for products and services, based on their past hires, defaults, duration of previous contracts, location, usage data, and connected devices used with the energy network. The ADM could potentially lead to a denial of service and follow-up advertising. According to the DPA, the privacy notice was difficult to understand without specialist knowledge. Users would likely be unaware of their right not to give consent or to withdraw it. The fact that much of the processing was not actually carried out was irrelevant.

Foodinho

In July 2021, the Italian DPA fined on-demand food delivery company Foodinho Srl €2.6 million for ADM by the company's excellence system called Jarvis in relation to peak time job allocation (*www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/9677611*). Foodinho attempted to rely on the exemption in Article 22(2)(a), which provides that Article 22(1) will not apply if the ADM was necessary for performing a contract.

The DPA held that Foodinho had failed to provide meaningful information about the logic involved in the ADM and the impact of outcomes, and had not implemented safeguards such as measures to guarantee the accuracy and correctness of the outputs, and individuals' rights to contest the decision, obtain human intervention and express their views.

Under the draft Digital Labour Platforms Directive, workers will be able to demand an explanation in writing, with the possibility of discussing and clarifying the facts, circumstances and reasons for ADM with a human operative. Appeals will have to be resolved within a week and incorrect decisions will have to be reversed or remedied with adequate compensation.

The DPA also noted the lack of a data protection impact assessment, access control measures, retention periods and AI oversight measures. It ordered Foodinho to verify the accuracy and relevance of the data used by the AI system and to identify measures to prevent the improper or discriminatory use of user reviews.

Deliveroo

In July 2021, the Italian DPA also fined Deliveroo €2.5 million for ADM by its algorithmic system called Frank in relation to job allocation (*www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/9685994*). The DPA rejected Deliveroo's argument that the algorithm merely caused a minor inconvenience to riders by potentially limiting jobs at their preferred times and place, and that Article 22 did not apply.

Customers can see a delivery time estimate, which is set before a job is assigned. Each order acceptance, arrival at the restaurant, food collection, arrival at the customer and delivery is tracked to feed the predictive analytics model for delivery time estimation. Drivers are scored on their reliability based on their performance, cancellations and availability. Geolocation data, mode of transport, and customer and restaurant location were also evaluated for job allocation. The DPA held that ADM based on this data could amplify or contribute to gender bias. The privacy notice rather vaguely explained that data was processed to guide the algorithms to make the most effective and accurate decisions, and to determine priority access level to bookings. The fact that the customer care team could track the live geolocation of each rider was considered a security failing.

Under the draft Digital Labour Platforms Directive, organisations will have to disclose key features of automated monitoring, including information about which specific work performance acts are monitored and evaluated. Systems will not be allowed to process personal data concerning platform workers that are not intrinsically connected to and strictly necessary for the performance of their contract, such as private conversations, health, emotional state and any data relating to the workers' time off.

Clearview AI

On 29 November 2021, the Information Commissioner's Office (ICO) announced that it intended to impose a £17 million fine on facial recognition company Clearview AI Inc for scraping images from the internet and using ADM to maintain a database of ten billion images (*https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2021/11/ico-issues-provisional-view-to-fine-clearview-ai-inc-over-17-million/*). Users can upload images of an individual's face, match it to photos of that person's

face collected from the internet and find out where those photos appeared. The images are scraped without the individuals' knowledge. The service was trialled for free by UK law enforcement agencies.

UK DATA REFORM

As part of its AI strategy, the government wants to tap into the huge potential for the linking and reusing of datasets across organisations, domains and sectors to allow AI to flourish (*see News brief "Artificial intelligence strategy: making the UK a global AI superpower", www.practicallaw.com/w-033-1152*). AI requires various sets of data, including training, input, output and baseline data, each of which highlights different data protection issues.

Training data

The government's data protection proposals promise to establish the freedom to experiment where it does not cause harm (*see News brief "Data protection reform: setting the course for a new direction", www.practicallaw.com/w-032-7584*). The government intends to liberalise access to data used to train AI systems under new provisions about lawful basis, transparency and anonymisation.

The proposals aimed at ensuring that research can be carried out without unnecessary recourse to consent include:

- A clear definition of research.
- A new lawful basis of data processing for research.
- Research purposes being explicitly declared "compatible further processing".
- Permission to process data for incompatible further purposes where necessary for important public interests.
- Allowing valid consent to be given for unknown future research.
- Private companies that undertake research for a public body being able to rely on the public task lawful basis for data processing.

Under the proposals, the legitimate interest for research would be predetermined without the need to conduct the balancing test between operational need and individuals' rights. However, in its 6 October 2021 response to proposals, the ICO pointed out that they do not remove the need for the balancing test; they instead move the responsibility for it from organisations to the government (*https://ico.org.uk/media/about-the-ico/consultation-responses/4018588/dcms-consultation-response-20211006.pdf*).

The government is also proposing that the transparency burden in relation to data sharing, retention and other aspects of research be reduced by clarifying that the disproportionate effort exemption under Article 14(5)(b) of the UK GDPR may apply in these scenarios. Under this narrow exemption, controllers that process data collected indirectly do not need to provide a privacy notice to the data subjects where this would cause disproportionate effort.

The government also proposes to put the motivated intruder test on a statutory footing. The motivated intruder test enables organisations to check whether there is a risk of individuals being identified from anonymised data by a motivated intruder. This would mean that personal data held by one controller would officially be considered anonymous data in the hands of another controller if it were impossible for that controller to identify the individuals without spending unreasonable time, effort or resources. However, the ICO has recently not shied away from elaborating on the complexities of this test in its new draft anonymisation guidance published in October 2021 (*https://ico.org.uk/media/about-the-ico/documents/4018606/chapter-2-anonymisation-draft.pdf*).

A significant development in the training of algorithms is the International Organization for Standardization's (ISO) new ISO/IEC TR 24027:2021 technical standard to address bias in relation to AI-aided decision making. The proposed holistic approach by the technical standard is underpinned by a management system standard that will consider the entire AI ecosystem with a view to continuously improve the specific AI system. The technical standard is intended to help create a system of checks and balances that ensures the ethical adoption of AI. The 2020 standard, ISO/IEC TR 24028:2020, tackles trustworthiness with a focus on transparency, explainability and security. These standards are expected to enable an AI-ready culture and fuel digital transformation.

Drawing on the successes of assurance frameworks in the information security industry, the government's roadmap to an effective AI assurance ecosystem, published on 8 December 2021, sets out plans to standardise and professionalise the AI assurance industry over the next five years (*www.gov.uk/government/publications/the-roadmap-to-an-effective-ai-assurance-ecosystem/the-roadmap-to-an-effective-ai-assurance-ecosystem*). AI assurance and certification will help to build trust and confidence in AI, which is essential for its adoption. Following on from the ICO's guidance on the AI auditing framework, the government's non-legislative initiatives will include an AI assurance guide from the Centre for Data Ethics and Innovation and a white paper on the governance of AI systems from the Office for AI, which is also expected to establish an AI standards hub (*https://cdeiuk.github.io/ai-assurance-guide/*) (*see News brief "New ICO guidance on AI: privacy by design saves retrofitting later", www.practicallaw.com/w-027-1236*).

Input data

The government wants to improve transparency to boost trust in AI. In particular, individuals want to understand how their input data will be handled and what input data may lead to which outcomes. Offering or denying public or private services on the basis of AI decisions will have significant effects on individuals.

In addition, the use of inferred data can lead to biased decisions. Users want to understand what inferences will be drawn about them and why; for example, the impact of belonging to an algorithmically defined group, such as people living in certain postcodes. The recent Deliveroo and Foodinho fines in Italy highlight the need for transparency when AI is relied on to make instantaneous decisions about an individual's eligibility (*see "Article 22" above*).

On 29 November 2021, the government released its Algorithmic Transparency Standard to ensure that the use of algorithmic tools to support decisions in the public sector are described in a complete, open, understandable, easily accessible and free format (*www.gov.uk/government/collections/algorithmic-transparency-standard*). It is intended to help people who use, regulate or are affected by the results of algorithmic-assisted decisions. The standard applies not just to ADM captured by Article 22 but any algorithmic tools used to support decisions,

regardless of human intervention and their legal, economic or similar impact on the individual (*see box "Algorithmic Transparency Standard"*).

In addition, on 24 November 2021, the government published the Product Security and Telecommunications Infrastructure (PSTI) Bill, which aims to ensure greater security in smart devices. The legislation would impose obligations on manufacturers, importers and distributors, and create an enforcement regime with civil and criminal sanctions aimed at preventing insecure products being made available on the UK market.

The PSTI Bill responds to predictions that there could be up to 50 billion connectable products worldwide by 2030, incidents such as the hacking of private security cameras where footage ended up on adult websites, and the potential physical harm that could be caused by smart heating elements or locks. The legislation would impose transparency about expected minimum security update support periods, a ban on default passwords and the requirement to implement a vulnerability disclosure policy to ensure that security researchers regularly warn manufacturers of security flaws.

Output data

The concept of fairness under the GDPR is said to include expectation fairness, fairness of process and outcome fairness. The government wishes to define fairness under the UK GDPR in narrower terms to boost experimental innovation. The ICO has already said that it is deeply concerned about any interference with fairness as the central principle in how individuals' data is processed. Indeed, a transcendent concept like fairness is arguably best suited to take on the growing complexities and unforeseen risks of AI.

While the government wants to remove Article 22 and fall back on the concept of legitimate interest to protect individuals' interests, the ICO is proposing to expand Article 22 to all ADM. Currently, Article 22 can be circumvented by introducing a human into the decision-making process, as seen in the Dutch *Uber* decisions (*see "Uber and Ola" above*). The consequence of this is that the controller will not have to provide meaningful information about the logic involved in its ADM.

However, it could be argued that, regardless of Article 22 being triggered, in order to establish legitimate interest in a valid way it will likely be necessary to provide fairly detailed information about the design, intended application, role in the decision-making process, risks and mitigations of AI. The explainability of ADM remains a challenge for controllers and, as seen in the Italian DPA's fining of Deliveroo, a vague and concise description will not suffice (*see "Deliveroo" above*). The government's new Algorithmic Transparency Standard may prove persuasive in this area of law.

AI oversight mechanisms are becoming an implicit requirement without which fairness, and therefore the lawfulness, of ADM cannot be achieved. If an AI technology provider does not have an algorithmic audit team with a track record of successful interventions, alarm bells should be ringing. AI must be scrutinised even if the metrics suggest positive results because the fast-paced and large-scale nature of AI and ADM means that, without active AI oversight, significant learning opportunities can be missed. Any AI provider that cannot present a data ethics policy would be best avoided. Without oversight reporting from the AI provider, there is unlikely to be compliance with the GDPR.

Baseline data

Baseline data is necessary for the benchmarking of AI results, preventing bias and ensuring the fairness of AI outputs. The government is proposing to introduce a new condition within Schedule 1 to the Data Protection Act 2018 to address the processing of sensitive personal data for bias monitoring, detection and correction in relation to AI systems. In addition, for data processing intended to safeguard important public interests, the government proposes to support data intermediaries in becoming trusted custodians of data that can be collected, shared or pooled.

Algorithmic Transparency Standard

The pilot version of the Algorithmic Transparency Standard consists of guidance, a transparency template and the algorithmic transparency data standard. Public authorities will submit the template to the government for formatting and publication. The template includes information about:

- How the algorithmic tool works.
- How the tool is incorporated into and affects the decision-making process, including how humans have oversight of the tool.
- The problem that the tool is intended to solve and how it solves that problem.
- The justification or rationale for using the tool, including what the tool is and is not intended for, benefits of the tools, and non-algorithmic alternatives considered.
- How people can find out more about the tool, ask a question, the appeal and review process.
- Who owns and has responsibility for the tool.
- The tool's technical specifications, including the type of model, volume of use, development phase and the system architecture.
- The datasets used to train the model and the datasets that the model will be used on.
- Impact assessments that have been carried out, such as data protection impact assessments, algorithmic impact assessments, ethical assessments and equality impact assessments.
- Common risks for the tool and mitigations.

LOOKING AHEAD

It is clear that although AI will bring value and benefits to mankind, this may be accompanied by geopolitical tensions, oppression and, in the worst case, extinction. AI will certainly offer a rare opportunity for humanity to self-reflect.

AI ethics is an essential ingredient. Without it, humanity could one day wake up to a reality where man serves AI instead of harnessing it. Adopting global principles for AI ethics is essential and the development of hostile AI must be stopped now to avoid catastrophic consequences for mankind. The international community has embarked on a path to develop global ethics, yet consensus remains distant. With a bit of optimism, it could be foreseen that, in the future, AI ethics will reflect on mankind and make society reconsider its values for the better.

When it comes to data, which is essential for the development of AI, the government is unlikely to be able to compete with tech hubs in jurisdictions that are not signatories to, or have failed to ratify, conventions on international human rights. However, the government's proposal to make it easier to carry out research without needing unnecessary consent is an important one and may help the UK to continue participating in the AI debate and influencing its future.

At the same time, Article 22 should probably not be removed from the UK GDPR but, as the ICO suggests, expanded instead. In the last two decades, the concept of legitimate interest has failed to protect people from the data needs of the adtech sector. If this is the government's best alternative to Article 22, it will equally fail to provide protection from the risks of AI. However, this time, the effects of data processing may not be limited to influencing our consumer behaviour. Further laws to safeguard humans from AI are in order, including a ban on certain technologies, such as remote facial recognition, as suggested in the UN High Commissioner for Human Rights' report on the right to privacy in the digital age (*see News brief "Facial recognition technology: the risks unfold", www.practicallaw.com/w-033-4793*). The examples of the EU's Artificial Intelligence Act and the draft Digital Labour Platforms Directive show merit.

Organisations today have the ability to explain their ADM in greater detail than ever before. With ISO standardisation on AI management and oversight, operating compliant AI is within reach. The direct liability of manufacturers and distributors of AI technologies under the PSTI Bill will contribute to a safer AI-powered future.

Alexander Dittel is a partner in the technology practice at Wedlake Bell LLP.

Related information

This article is at practicallaw.com/w-033-8467

Other links from uk.practicallaw.com/

Topics

Data protection offences	topic/1-607-9647
GDPR and data protection reform	topic/7-616-6199
Human rights	topic/6-103-1253
Information technology	topic/5-103-2074
Surveillance: data protection	topic/7-616-6222
Technology: data protection	topic/8-616-6207

Practice notes

Assessing a machine learning model for non-data scientists	w-024-5946
Data subject rights (UK)	w-024-3178
Demystifying artificial intelligence	w-008-5369
Legal aspects of artificial intelligence	w-018-2338
Overview of UK GDPR	w-013-3757
Processor obligations under UK GDPR	w-025-2861
UK GDPR and DPA 2018: exemptions	w-014-6104
UK GDPR and DPA 2018: profiling and automated decision-making	w-014-3599

Previous articles

GDPR enforcement: a changed landscape (2021)	w-030-5470
The adtech challenge: thriving in an e-commerce world (2021)	w-032-9223
AI and data protection: balancing tensions (2019)	w-020-9713
Big data and competition law: with great opportunities comes great risks (2019)	w-020-7091
Challenges in the consumer sector: transformative technology (2019)	w-020-3706
Data assets: protecting and driving value in a digital age (2019)	w-019-8276
Data protection: privacy by (re)design (2019)	w-018-6087
E-Privacy Regulation: developing slowly (2019)	w-020-8272
GDPR one year on: taking stock (2019)	w-020-0982
Algorithms, apps and AI: the next frontier in discrimination law (2018)	w-013-8054
Artificial intelligence: navigating the IP challenges (2018)	w-015-2044
Data use: protecting a critical resource (2018)	w-012-5424

For subscription enquiries to Practical Law web materials please call +44 0345 600 9355