Journal of Responsible Technology 4 (2020) 100005

Contents lists available at ScienceDirect

Journal of Responsible Technology

journal homepage: www.elsevier.com/locate/jrt

Legal and human rights issues of AI: Gaps, challenges and vulnerabilities

Rowena Rodrigues*

Trilateral Research Ltd.One Knightsbridge Green (5th Floor), London SW1x7QA UK

ARTICLE INFO

Keywords:
Artificial intelligence
AI
Legal issues
Human rights
Vulnerability

ABSTRACT

This article focusses on legal and human rights issues of artificial intelligence (AI) being discussed and debated, how they are being addressed, gaps and challenges, and affected human rights principles. Such issues include: algorithmic transparency, cybersecurity vulnerabilities, unfairness, bias and discrimination, lack of contestability, legal personhood issues, intellectual property issues, adverse effects on workers, privacy and data protection issues, liability for damage and lack of accountability. The article uses the frame of 'vulnerability' to consolidate the understanding of critical areas of concern and guide risk and impact mitigation efforts to protect human well-being. While recognising the good work carried out in the AI law space, and acknowledging this area needs constant evaluation and agility in approach, this article advances the discussion, which is important given the gravity of the impacts of AI technologies, particularly on vulnerable individuals and groups, and their human rights.

Introduction

Artificial intelligence (AI)[1] is everywhere (Boden 2016) and its development, deployment and use is moving forward rapidly and contributing to the global economy (McKinsey 2019; PwC 2017). AI has many benefits (e.g., improvements in creativity, services, safety, lifestyles, helping solve problems) and yet at the same time, raises many anxieties and concerns (adverse impacts on human autonomy, privacy, and fundamental rights and freedoms) (OECD 2019).

The legal discourse on the legal and human rights issues of artificial intelligence (AI) is established, with many a detailed legal analysis of specific individual issues (as outlined in Sections 3 and 4 in this article). But, this field is a regulatory moving target and there is a need for an exploratory, bird's eye and looking at the breadth of issues, curated in a single place. Critically missing also is a greater discussion and mapping of vulnerability to such issues. This article fills this gap based on research carried out in the EU-funded Horizon 2020 SIENNA project[2]. The article's main research questions are: What are the legal and human rights issues related to AI? (How) are they being addressed? What are the gaps and challenges and how can we address vulnerability and foster resilience in this context?

Structure, approach, method and scope

After a quick round-up of the coverage of legal and human rights issues (Section 3), this article outlines specific legal issues being discussed in relation to AI (Section 4), solutions that have been proposed/how they are being addressed, gaps and challenges, and affected human rights principles (Section 5). It maps the legal issues to core international human rights treaties and provides examples (global to regional) of corresponding human rights principles that might be affected. More vitally, it discusses the legal issues using the frame of 'vulnerability' (Section 6) to help consolidate better the identification of what are critical areas of concern and help guide AI risk and impact mitigation efforts to protect human and societal well-being. While recognising the good work already being carried out in the AI law space (as evident in the literature identified in this article), this consolidated analysis of issues hopes to further provide insights and add to the much-required need for further and sustained discussions on this topic, given AI's increasingly widespread deployment and use and the gravity of its impacts on individuals and their human rights.

There are a number of legal issues and human rights challenges related to AI.[3] Section 4 presents a panoramic, non-exhaustive overview

* Corresponding author.
E-mail address: rowena.rodrigues@trilateralresearch.com

[1] Referring here to "systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals". European Commission (2018b).

[2] https://www.sienna-project.eu

[3] We also note the connection of AI and other technologies e.g., robotics; they might converge and be interconnected (e.g., artificially intelligent robots or software robots) and present similar challenges, but they are also each, distinct technologies, and serve different purposes. The focus here exclusively on AI which was also a great help to disentangle issues. However, many AI and robotics issues are inter-related (e.g., transparency, fairness, accountability) and might not operate in silos.

https://doi.org/10.1016/j.jrt.2020.100005
Received 17 August 2020; Accepted 27 September 2020
Available online 16 October 2020

of such issues[4] and challenges. The identification of issues was carried out using a desktop literature review (in two phases: preliminary research in 2018 as part the SIENNA project Rodrigues (2019) and updated in July 2019 during the development of this article). The keywords 'legal/human rights issues+AI/artificial intelligence/machine learning' were used to identify issues covered in legal academic and practitioner journals and books and legal policy studies from the last five to ten years (as cited in the article) supplemented by databases such as SSRN and Google Scholar, to identify issues high impact. The references that came to the forefront in our search were scanned further, as possible, for any other relevant unidentified issues. The inclusion of issues was conditioned by their coverage and/or prevalence in existing legal and policy literature, impact on societal values and life, and controversiality. One limitation was that this was a study limited by time and to research available in English. Furthermore, while each of these issues could be analysed in greater depth individually (e.g., looking into specific legal provisions that are applicable), this is outside the scope of study here and in many cases has been/is being carried out by other scholars.

For the mapping of legal issues to principles in international human rights treaties, we scanned the core international human rights instruments for coverage of such issues. These included, E.g., International Covenant on Civil and Political Rights (ICCPR), International Covenant on Economic, Social and Cultural Rights (ICESCR), Universal Declaration of Human Rights (UDHR), Charter of the United Nations, Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May be Deemed to be Excessively Injurious or to Have Indiscriminate Effects (and Protocols), and the European Convention for the Protection of Human Rights and Fundamental Freedoms etc.

For the mapping of identified AI legal issues to the most vulnerable groups and the factors that determine and/or facilitate vulnerability, we followed a law-in-context approach. The vulnerable groups and factors that determine vulnerability were identified and determined by scanning the literature reviewed in the issue identification, supplemented by an online search for further examples. The table presented is non-exhaustive and will change with examination in different contexts.

Coverage of legal and human rights issues

International coverage of legal and human rights issues is evident in policy documents of United Nations (2019); OECD (2019); Council of Europe (2017; 2018; 2019), the European Parliament (2017; 2018a; 2018b; 2019, 2020a, 2020b, 2020c), the European Commission (2018a, 2018b, 2020), European Commission for the efficiency of justice (CEPEJ) (2018), and the European Data Protection Supervisor (2016).

Academic and civil society (Access Now 2018; Privacy International and Article 19 2018) coverage of legal issues pertaining to AI sometimes are broad and cover a variety of risks and challenges. At other times, these cover very specific issues. Some analyses are domain-specific. E.g., focussing on healthcare (Price 2017), defence (Cummings 2017), transport (Niestadt et al 2019). Some of these include coverage of issues related to legal personality, intellectual property (Schönberger 2018), algorithmic bias, discrimination, unfairness (Smith 2017; Danks and London 2017; Courtland 2018; Hacker 2018), labour protection (De Stefano 2019), privacy and data protection (Wachter and Mittelstadt 2019), cybersecurity (Tschider 2018), access to justice (Raymond 2013), algorithmic transparency (Lepri 2018; Coglianese and Lehr 2018; Bodo et al 2018), liability for harms (Vladeck 2014), accountability (Liu et al 2019), and surveillance (Aloisi and Gramano 2019; Feldstein 2019).

The media coverage of AI legal issues has ranged from the broad (Dizikes 2019) to more specific - covering aspects such as liability (Mitchell 2019), fairness in decision-making (Niiler 2019), bias (Marr 2019), privacy (Lindsey 2018), accountability (Coldewey 2018). Issues of privacy/data protection (Meyer 2018; Williams 2019; Forbes Insights Team 2019; Lohr 2019) and bias (Dave 2018), in particular, have received significant publicity.

Legal and human rights issues of AI

This section briefly examines each issue, its significance, solutions that have been proposed (or how is it being addressed) and the related gaps and challenges. This is a limited analysis (other research has analysed and critically discussed each of these issues in detail; the intent here is to provide a panoramic, updated overview and make it useful for future research).

Of the ten issues presented below, some relate to the *design and nature* of AI itself (these are covered first), others are issues connected to the implementation and use of AI (though often, the design of AI itself lends itself to causing or facilitating implementation and use issues). The issues are sometimes cross-domain, i.e., could manifest in one or more sector/field of application. Many of these issues are common to all technology (e.g., privacy/data protection); many are inter-related (e.g., transparency, fairness, accountability) and might not operate in silo. However, the ability of AI to amplify and/or facilitate their adverse effects must not be underestimated at any time.

Lack of algorithmic transparency

The issue and its significance

The lack of algorithmic transparency (Bodo et al 2018; Coglianese & Lehr 2018; Lepri et al 2018) is a significant issue that is at the forefront of legal discussions on AI (EDPS2016; Pasquale 2015). Cath (2018) highlights given the proliferation of AI in high-risk areas, that "pressure is mounting to design and govern AI to be accountable, fair and transparent." The lack of algorithmic transparency is problematic; Desai and Kroll (2017) highlight why, using examples of people who were denied jobs, refused loans, were put on no-fly lists or denied benefits without knowing "why that happened other than the decision was processed through some software". Information about the functionality of algorithms is often intentionally poorly accessible" (Mittelstadt et al 2016) and this exacerbates the problem.

Solutions proposed/how it is being addressed

An EU Parliament STOA study (2019) outlined various policy options to govern algorithmic transparency and accountability, based on an analysis of the social, technical and regulatory challenges; each option addresses different aspect of algorithmic transparency and accountability: 1. awareness raising: education, watchdogs and whistle blowers; 2. accountability in public-sector use of algorithmic decision-making; 3. regulatory oversight and legal liability; and 4. global coordination for algorithmic governance. More specific solutions mooted to promote algorithmic transparency include algorithmic impact assessments (Reisman, et al 2018; Government of Canada, undated), an algorithmic transparency standard (IEEE P7001:Transparency of Autonomous Systems), counterfactual explanations, local interpretable model-agnostic explanations (LIME) (Ribeiro, Singh, Guestrin 2016) etc.

Gaps and challenges

Transparency has its limitations and is often viewed as inadequate and limited (Ananny and Crawford 2018). For example, as Vaccaro and Karahalios (undated) point out, "Even when machine learning decisions can be explained, decision-subjects may not agree with the outcome". Some of solutions proposed above, e.g., algorithmic impact assessments, though extremely valuable, are relatively new and still a work in progress so cannot be fully evaluated for their effectiveness at this stage. This is definitely an area for future research and evaluation.

[4] The order of presentation of the issues does not reflect a hierarchy.

Cyber security vulnerabilities

The issue and its significance

A RAND perspectives report Osoba and Welser (2017) highlights various security issues related to AI, for example, fully automated decision-making leading to costly errors and fatalities; the use of AI weapons without human mediation; issues related to AI vulnerabilities in cyber security; how the application of AI to surveillance or cyber security for national security opens a new attack vector based on 'data diet vulnerability'; the use of network intervention methods by foreign-deployed AI; larger scale and more strategic version of current advanced targeting of political messages on social media etc. The report Osoba and Welser (2017) also identifies domestic security-related issues, for example, (growing) deployment of artificial agents for the surveillance of civilians by governments (e.g., predictive policing algorithms). These have been called out for their potential to adversely impact fundamental citizens' rights Couchman (2019). Such issues are significant as they lay open critical infrastructures to harms with severe impacts on society and individuals, posing a threat to life and human security and access to resources. Cyber security vulnerabilities also pose a significant threat as they are often hidden and revealed only to late (after the damage is caused).

Solutions proposed/how it is being addressed

Various strategies and tools are being used or proposed to address this issue. E.g., putting in place good protection and recovery mechanisms; considering and addressing vulnerabilities in the design process; engaging human analysts in critical decision-making; using risk management programmes; and software upgrades Fralick (2019).

Gaps and challenges

Effectively addressing such issues requires proactive and responsive use of cybersecurity policies, mechanisms and tools by developers and users at all stages – design and implementation and use. But this is often not the case in practice and is a real challenge As a SHERPA report outlines, "*When designing systems that use machine learning models, engineers should carefully consider their choice of a particular architecture, based on understanding of potential attacks and on clear, reasoned trade- off decisions between model complexity, explainability, and robustness*" (Patel et al, 2019).

"在設計使用機器學習模型的系統時，工程師應根據對潛在攻擊的理解以及模型複雜性、可解釋性和穩健性之間的明確、合理的權衡決定，仔細考慮他們對特定架構的選擇"

Unfairness, bias and discrimination

The issue and its significance

Unfairness (Smith 2017), bias (Courtland 2018) and discrimination (Smith 2017) repeatedly pop up as issues and have been identified as a major challenge (Hacker 2018) related to the use of algorithms and automated decision-making systems, e.g., to make decisions related to health (Danks & London 2017), employment, credit, criminal justice (Berk 2019), and insurance. In August 2020, protests were made and legal challenges are expected over the use of a controversial exams algorithm used to assign grades to GCSE students in England (Ferguson & Savage 2020).

A focus paper from the EU Agency for Fundamental Rights (FRA 2018) outlines the potential for discrimination against individuals via algorithms, and states that "the principle of non-discrimination, as enshrined in Article 21 of the Charter of Fundamental Rights of the European Union, needs to be taken into account when applying algorithms to everyday life" (FRA 2018). It cites examples with potential for discrimination: automated selection of candidates for job interviews, use of risk scores in creditworthiness or in trials. An European Parliament report on the fundamental rights implications of big data: privacy, data protection, non-discrimination, security and law-enforcement, European Parliament (2017) stressed that "*because of the data sets and algorithmic systems used when making assessments and predictions at the different stages of data processing, big data may result not only in infringements of the fundamental rights of individuals, but also in differential treatment of and indirect discrimination against groups of people with similar characteristics, particularly with regard to fairness and equality of opportunities for access to education and employment, when recruiting or assessing individuals or when determining the new consumer habits of social media users*" European Parliament (2017). The report called on the European Commission, the Member States and data protection authorities "*to identify and take any possible measures to minimise algorithmic discrimination and bias and to develop a strong and common ethical framework for the transparent processing of personal data and automated decision-making that may guide data usage and the ongoing enforcement of Union law*"European Parliament (2017).

Solutions proposed/how it is being addressed

Various proposals have been made to address such issues. For example, conducting regular assessments into the representativeness of data sets and whether they are affected by biased elements European Parliament (2017), making technological or algorithmic adjustments to compensate for problematic bias (Danks & London 2017), humans-in-the-loop (Berendt, Preibusch 2017) and making algorithms open. Schemes to certify that algorithmic decision systems do not exhibit unjustified bias are also being developed. The IEEE P7003 Standard for Algorithmic Bias Considerations is one IEEE ethics-related standards (under development as part of the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems) - it aims to provide individuals or organisations creating algorithmic systems with development framework to avoid unintended, unjustified and inappropriately differential outcomes for users. There are also open source toolkits, e.g., the AI Fairness 360 Open Source Toolkit that helps users to examine, report, and mitigate discrimination and bias in machine learning models throughout the AI application lifecycle. It uses 70 fairness metrics and 10 state-of-the-art bias mitigation algorithms developed by the research community.

Gaps and challenges. While the law clearly regulates and protects against discriminatory behaviour, it is suggested it falls short. A Council of Europe (2018) study outlines that the law leaves shortfalls where it does not extend to address what is not expressly protected against discrimination by law, or where new classes of differentiation are created and lead to biased and discriminatory effects. Humans-in-the-loop approaches might face tensions as to where, and in which cases they should be applied (sometimes it might be better to not to have, or impossible to have humans in the loop, e.g., where there might be scope for human error or stupidity that leads to serious or irreversible consequences). Other gaps include whether the use of human-in-the-loop is adequately signified in the technologies that use them. Making algorithms open does not mean they will become more understandable to people, also there is the issue of the exposure or discoverability of private data that brings its own concerns House of Commons (2018). Algorithmic auditing to be effective requires a holistic, interdisciplinary, scientifically-grounded and ethically-informed approach (Guszca et al 2018). While the technical solutions proposed thus, are good steps forward there have been many calls to pay greater regulatory, policy and ethical attention to *fairness*, especially in terms of protection of vulnerable and marginalised populations Raji & Buolamwini (2019).

Lack of contestability

The issue and its significance

European Union data protection law gives individuals rights to challenge and request a review of automated decision-making that significantly affects their rights or legitimate interests (GDPR 2016/679). Data subjects have the right to object, on grounds relating to their particular situation, at any time to the processing of personal data concerning them which is based on tasks carried out in public interests or legitimate interests. Further, per Article 22(3) GDPR, data controllers must implement suitable measures to safeguard a data subject's rights and freedoms

and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express their point of view and to contest the decision. But Hildebrandt (2016) underlines how "the opacity of ML systems may reduce both the accountability of their 'owners' and the contestability of their decisions". Edwards and Veale (2017) highlight, the lack of contestability - in relation to algorithmic systems, i.e., the "lack of an obvious means to challenge them when they produce unexpected, damaging, unfair or discriminatory results". Bayamlıoğlu (2018) states that "*a satisfactory standard of contestability will be imperative in case of threat to individual dignity and fundamental rights" and "the 'human element' of judgment is, at least for some types of decisions, an irreducible aspect of legitimacy in that reviewability and contestability are seen as concomitant of the rule of law and thus, crucial prerequisites of democratic governance*".

Solutions proposed/how it is being addressed

Contestability by design has been proposed as an approach to better protect the rights of decisions based solely on automated processing as a requirement at each stage of an artificial intelligence system's lifecycle Almada (2019).

Gaps and challenges

As Roig (2017) argues, the "*general safeguards – specific information to the data subject; the right to obtain human intervention; the right to express his or her point of view; the right to obtain an explanation of the decision reached; and the right to challenge the decision – may not work in the case of data analysis-based automated processing*". Further, that it "*will be difficult to contest an automatic decision without a clear explanation of the decision reached. To challenge such an automatic data-based decision, only a multidisciplinary team with data analysts will be able to detect false positives and discriminations*" Roig (2017). So, this is an issue that needs to be further addressed at many different levels (design, development and use).

Legal personhood issues

The issue and its significance

There is ongoing debate about whether AI (and/or robotics systems) "fit within existing legal categories or whether a new category should be created, with its own specific features and implications". (European Parliament Resolution 16 February 2017). This is not just a legal, but a politically-charged issue Burri (2017). Čerka et al (2017), ask whether AI systems can be deemed subjects of law. The High-Level Expert Group on Artificial Intelligence (AI HLEG) has specifically urged "policy-makers to refrain from establishing legal personality for AI systems or robots" outlining that this is "fundamentally inconsistent with the principle of human agency, accountability and responsibility" and poses a "significant moral hazard" (AI HLEG 2019). Yet, others such as Turner (2019) suggest that "legal personality for AI could be justified as an elegant solution to (i) pragmatic concerns arising from the difficulties of assigning responsibility for AI and/or (ii) in order to support AI's moral rights, if any". Jaynes (2020) assumes that in the future artificial entities will be granted citizenship and discusses the jurisprudence and issues pertaining to non-biological intelligence that are important to consider. In the EU, at least, however the general caution to avoid creating new legal personality for AI systems has been echoed manifold (Siemaszko, Rodrigues, Slokenberga, 2020; Bryson, Diamantis, and Grant 2017)

Solutions proposed/how it is being addressed

There has not been a significant breakthrough in addressing legal personhood issues for AI at the international, EU or national level. While this issue has been raised (and will continue to be at the forefront of legal debates for the near future), international or even regional-level agreement Delcker (2018) on this (i.e., whether legal personhood should be offered to AI systems/robots and the form this should take) might be difficult or near impossible to achieve (given the political nature and sensitivity of the issue). Further, such issues are largely regulated at the national level.

Gaps and challenges

Brożek and Jakubiec (2017) investigated the issue of legal responsibility of autonomous machines and argue that "autonomous machines cannot be granted the status of legal agents." Bryson, Diamantis, and Grant (2017) consider conferring legal personhood on purely synthetic entities will become a very real legal possibility but think such "legislative action would be morally unnecessary and legally troublesome". In their review of the utility and history of legal fictions of personhood and after discussing the salient precedents where such fictions resulted in abuse or incoherence, they argue that, "While AI legal personhood may have some emotional or economic appeal, so do many superficially desirable hazards against which the law protects us" Bryson, Diamantis, and Grant (2017).

Intellectual property issues

The issue and its significance

Intellectual property rights are part of the Universal Declaration of Human Rights (UDHR, Article 27), the International Covenant on Economic, Social and Cultural Rights (ICESCR, Article 15), the International Covenant on Civil and Political Rights (ICCPR, Article 19) and the Vienna Declaration and Programme of Action (VDPA) 1993. Such rights they have a "human rights character" and "have become contextualised in diverse policy areas"WIPO (1998). AI raises various intellectual property issues, e.g., who owns AI generated/produced works or inventions? Should AI's inventions be considered prior art? Who owns the dataset from which an artificial intelligence must learn? Who should be liable for creativity and innovation generated by AI, if they impinge upon others' rights or other legal provisions? (CEIPI undated).

Solutions proposed/how it is being addressed

The law may provide a variety of solutions for the issues raised Rodrigues (2019). For example, in the UK, the law protects computer-generated literary, dramatic, musical or artistic works. There is no express legal provision on patentability of computer-generated works. The creator of the AI design owns such rights except if the work was commissioned or created during the course of employment. In this latter case, the rights belong to the employer or party that commissioned the AI work UK Copyright Service (2004). As a registered trade mark is personal property, unless an AI system was able to hold/have personal property, this right might not apply or be able to be enjoyed by the AI system.

Gaps and challenges

Many intellectual property rights issues have not been addressed and/or answered conclusively, and current regimes have been seen as "woefully inadequate to deal with the growing use of more and more intuitive artificial intelligence systems in the production of such works" Davies (2011). There is need further research and exploration especially as AI advances further and it becomes increasingly difficult to identify the creator. Talking Tech (2017).

Adverse effects on workers

The issue and its significance

The IBA Global Employment Institute report (2017) highlights the impact of AI and robotics on the workplace (seen a global concern). Some issues highlighted include: changes to the requirements for future employees, lowering in demand for workers, labour relations, creation of new job structures and new types of jobs, dismissal of employees, inequality in the 'new' job market, integration of untrained workers in the 'new' job market, labour relations (and its possible implications for union activities and collective bargaining aspects, challenges for employee representatives, changes in the structure of unions), health and safety issues, impact on working time, impact on remuneration (changes, pensions), social security issues etc. Significant is also

the potential loss of autonomy for workers Frontier Economics (2018). These issues have economic (e.g., poverty) and social consequences (e.g., homelessness, displacement, violence, despair) and significant human rights impact potential. They raise ethical issues and dilemmas that might not easily be resolved yet are critical to address.

Solutions proposed/how it is being addressed

Many measures or solutions are being or have been proposed to address this issue. These include retraining workers (UK House of Lords 2018) and re-focussing and adapting the education system. The Communication from the European Commission on Artificial Intelligence for Europe (2018), suggests the modernisation of education, at all levels, should be a priority for governments and that all Europeans should have every opportunity to acquire the skills they need. To manage the AI transformation, the Communication calls for providing support to workers whose jobs change or disappear – it suggests "national schemes will be essential for providing such up-skilling and training. (European Commission on Artificial Intelligence for Europe 2018). Social security systems will also require review and change.

Gaps and challenges

One report prepared for the Royal Society (2018) highlights gaps in the evidence base, particularly in relation to there being "*limited evidence on how AI is being used now and on how workers' tasks have changed where this has happened*", "*relatively little discussion of how existing institutions, policies, social responses are shaping and are likely to shape the evolution of AI and its adoption*" and "*little consideration of how international trade, mobility of capital and of AI researchers are shaping the development of AI and therefore its potential impact on work*" Frontier Economics (2018). While there is recognition of the widespread disruption that AI is, and might create in the workplace, not enough has been put in place at the policy and regulatory level to address concerns and put in place needed economic and educational policies and measures. At the employer-level too, while AI solutions are being widely deployed, it remains to be seen whether employers will adopt suitable strategies or due diligence checks to minimise any adverse impacts on their workforces and help them adapt or adjust to a changed workplace.

Privacy and data protection issues

The issue and its significance

Legal scholars and data protection enforcement authorities (CNIL 2017; ICO 2017) believe that AI (in addition to affecting other rights) poses huge privacy and data protection challenges Gardner (2016). These include informed consent, surveillance Brundage (2018)[5], infringement of data protection rights of individuals, e.g., right of access to personal data, right to prevent processing likely to cause damage or distress, right not to be subject to a decision based solely on automated processing etc.). Wachter & Mittelstadt (2019) highlight concerns about algorithmic accountability and underline that "*individuals are granted little control and oversight over how their personal data is used to draw inferences about them*" and call for a new data protection 'right to reasonable inferences' to close the accountability gap posed 'high risk inferences' – i.e., inferences that are privacy invasive or reputation damaging and have low verifiability in the sense of being predictive or opinion-based" Wachter & Mittelstadt (2019).

The EDPS, *Artificial Intelligence, Robotics, Privacy and Data Protection Background document for the 38th International Conference of Data Protection and Privacy Commissioners* 2016, highlighted the potential for increase in privacy implications and powerfulness of surveillance possibilities. The UK Information Commissioner's Office (ICO)'s discussion paper on *Big data, artificial intelligence, machine learning and data protection* (2017) examined the implications of big data, artificial intelligence (AI) and machine learning for data protection, highlights the intrusive nature of big data profiling and the challenges for transparency (due to the complexity of methods used in big data analysis) (ICO 2017).

Solutions proposed/how being addressed

Privacy and data protection law (particularly in the European Union) provides, at least in the letter of the law, good safeguards and protection for infringement of data subjects' rights. E.g., GDPR rights of data subjects to transparency, information and access (Article 15), rectification (Article 16) and erasure (Article 17), right to object to automated individual decision-making (Article 21) etc.

In relation to informed consent in the use of AI, transparency of potential harms relating to its use is strongly supported (Rigby 2019); developers should "*pay close attention to ethical and regulatory restrictions at each stage of data processing. Data provenance and consent for use and reuse are considered to be of particular importance*" (Vayena, Blasimme & Cohen 2018). In relation to surveillance, Brundage et al suggest secure multiparty computation (MPC) (which "refers to protocols that allow multiple parties to jointly compute functions, while keeping each party's input to the function private" (Brundage 2018). Other measures that are being used or proposed include the use of anonymisation, privacy notices, privacy impact assessment, privacy by design, use of ethical principles and auditable machine algorithms (ICO 2017).

Gaps and challenges

Privacy and data protection law does not address all AI issues. As pointed out, "*understanding and resolving the scope of data protection law and principles in the rapidly changing context of AI is not an easy task, but it is essential to avoid burdening AI with unnecessary regulatory requirements or with uncertainty about whether or not regulatory requirements apply*"(CIPL 2018). Privacy and data protection measures are only effective if they are used, properly applied, monitored and/or enforced. Also, e.g., as the European Data Protection Supervisor *Opinion 5/2018 Preliminary Opinion on privacy by design* points out, "*there is a limited uptake of commercial products and services fully embracing the concept of privacy by design and by default*". In some cases, the challenge is that the effectiveness of measures such as privacy/data protection impact assessments, privacy by design might fall flat (like closing the gate after the horse has bolted) given the core purpose of the AI system or technology by itself might conflict directly with societal values and fundamental rights.

Wachter and Mittelstadt (2019), argue that as the GDPR provides insufficient protection against sensitive inferences (Article 9) or remedies to challenge inferences or important decisions based on them (Article 22(3)), a new data protection right, the 'right to reasonable inferences', is needed to help close the accountability gap currently posed by *'high risk inferences'*. This would be useful particularly when this issue cannot or fails to be addressed via other means outlined above.

Liability for damage

The issue and its significance

The deployment and use of AI technologies can cause damage to persons and property. E.g., Gluyas and Day (2018) provide some examples – e.g., running over of pedestrians by driverless cars, crashing and damage caused by a partially operated drone, wrongful medical treatment diagnosis by an AI software programme. They further explain, "*As there are many parties involved in an AI system (data provider, designer, manufacturer, programmer, developer, user and AI system itself), liability is difficult to establish when something goes wrong and there are many factors to be taken into consideration…*" Gluyas & Day (2018).

Solutions proposed/how it is being addressed

Liability issues of AI could be addressed under the purview of civil or criminal liability. Kingston (2016) discusses AI and legal liability – both

[5] E.g., Brundage et al (2018) outline how "the use of AI to automate tasks involved in surveillance (e.g. analysing mass-collected data), persuasion (e.g. creating targeted propaganda), and deception (e.g. manipulating videos) may expand threats associated with privacy invasion and social manipulation".

whether criminal liability could ever apply, to whom it might apply, and, under civil law, whether an AI program is a product that is subject to product design legislation (product liability, e.g., in cases of design or manufacturing failures) or a service to which the tort of negligence applies.

Hallevy (2015) discusses the criminal liability of AI entities, i.e., responsibility for harm caused and explores whether an AI entity itself be criminally liable (beyond the criminal liability of the manufacturer, end-user or owner, and beyond their civil liability) and suggests that the imposition of criminal liability upon AI entities for committing intellectual property offenses is quite feasible and proposes solutions for sentencing AI entities. Liability issues could also be addressed under consumer protection law.

Rachum-Twaig (2020) proposes "*supplementary rules that, together with existing liability models, could provide better legal structures that fit AI-based robots. Such supplementary rules will function as quasi-safe harbors or predetermined levels of care. Meeting them would shift the burden back to current tort doctrines. Failing to meet such rules would lead to liability. Such safe harbors may include a monitoring duty, built-in emergency breaks, and ongoing support and patching duties*." Rachum-Twaig argues that "*these supplementary rules could be used as a basis for presumed negligence that complements the existing liability models*".

Gaps and challenges

In certain civil law jurisdictions, many liability issues are handled through strict liability. However, Bathee (2018) outlines "Strict liability is also a poor solution for the problem because if one cannot foresee the solutions an AI may reach or the effects it may have, one also cannot engage in conduct that strict liability is designed to incentivize, such as taking necessary precautions or calibrating the level of financial risk one is willing to tolerate". The European Commission *Expert Group on Liability and New Technologies* (2019) concluded in its review of existing liability regimes on emerging digital technologies, "*that the liability regimes in force in the Member States ensure at least basic protection of victims whose damage is caused by the operation of such new technologies. However, the specific characteristics of these technologies and their applications – including complexity, modification through updates or self- learning during operation, limited predictability, and vulnerability to cybersecurity threats – may make it more difficult to offer these victims a claim for compensation in all cases where this seems justified. It may also be the case that the allocation of liability is unfair or inefficient. To rectify this, certain adjustments need to be made to EU and national liability regimes*." In 2020, the European Commission published a Report on the safety and liability framework European Commission (2020). The European Parliament Legal Affairs (JURI) committee discussed in May 2020 a draft report on AI civil liability European Parliament (2020a).

Lack of accountability for harms

The issue and its significance

As outlined by the Assessment List for Trustworthy AI (ALTAI), accountability calls for mechanisms be put in place to ensure responsibility for the development, deployment and/or use of AI systems - risk management, identifying and mitigating risks in a transparent way that can be explained to and audited by third parties (AI HLEG 2020). As outlined by Dignum (2018), "*accountability in AI requires both the function of guiding action (by forming beliefs and making decisions), and the function of explanation (by placing decisions in a broader context and by classifying them along moral values)*". Some commentators suggest that "'accountability gap' is a worse problem than it might first seem" causing problems in three areas: causality, justice, and compensation Bartlett (2019). As a Privacy International and Article 19 (2018) report states, "Even when a potential harm is found, it can be difficult to ensure accountability for violations of those responsible."

Solutions proposed/how it is being addressed

Wachter, Mittelstadt, and Floridi (2017) suggest that "American and European policies now appear to be diverging on how to close current accountability gaps in AI". Legal accountability mechanisms for AI harms might take the form of a 'right to explanation' Edwards, Veale (2017), data protection and information and transparency safeguards, auditing, or other reporting obligations. Doshi-Velez et al (2017) review contexts in which explanation is currently required under the law and outline technical considerations that must be considered if it is desired that AI systems that could provide kinds of explanations that are currently required of humans.

Gaps and challenges

As Bartlett (2019) outlines, *"There is no perfect solution to AI accountability. One of the biggest risks with the proposal to hold developers responsible is a chilling effect on AI development. After all, AI developers are often small actors - individuals or small companies. Whether or not they are the most culpable when their creations cause harm, the practical nightmare of facing lawsuits every time their AI causes damage might reasonably make AI developers exceedingly wary of releasing their creations into the world (and their hedge fund investors might pause before reaching for their cheque books)*" Bartlett (2019). The right to explanation, as an accountability tool, has its challenges. As Wallace (2017) points out, "it is often not practical or even possible, to explain all decisions made by algorithms". Further, "the challenge of explaining an algorithmic decision comes not from the complexity of the algorithm, but the difficulty of giving meaning to the data it draws on"Wallace (2017). Edwards & Veale (2017) have argued extensively why a right to an explanation in the GDPR is unlikely to present a complete remedy to algorithmic harms (and might even lead to the creation of a transparency fallacy or be distracting). They suggest the law is restrictive, unclear, and even paradoxical concerning when any explanation-related right can be triggered. They further outline how "the legal conception of explanations as "meaningful information about the logic of processing" may not be provided by the kind of ML "explanations" computer scientists have developed, partially in response" Edwards & Veale (2017).

As one can see, there are a variety of legal issues pertaining to AI; some common problems of ICT technology in general - though facilitated or exacerbated by AI in some way, and other issues are novel and developing. All the issues will need to be kept in constant review to ensure that they are being appropriately addressed. We next examine the affected human rights principles.

Affected human rights principles

International human rights treaties lay down obligations which their signatories are bound to respect and fulfil; States must refrain from interfering with rights and take positive actions to fulfil their enjoyment. While, none of them currently explicitly apply or mention 'artificial intelligence/AI or machine learning', their broad and general scope would cover most of the issues and challenges identified. The table below maps the legal issues to human rights principles (drawn from the core international human rights treaties) that might be affected. In many cases, the affected principle is clear and obvious in others not so and needs to be drawn attention to.

Out of the affected human rights principles, widely prevalent in AI legal discussions, are the right to privacy and data protection (this is very prominent in Europe) and non-discrimination. Discussions also abound on the equality and access to justice. The remaining affected principles have been discussed but could benefit from much more airtime and future legal research.

Issues and vulnerability

It is not enough to simply outline the legal issues, gaps and challenges and the human rights principles AI implicates. Discussing these using the

frame of 'vulnerability' will valuably help consolidate the identification of critical areas of concern and guide AI risk and impact mitigation efforts to better protect human and societal well-being. It will also ensure that AI technologies advance human rights of everyone, and especially those most affected.

Vulnerability definitions are fragmented. Generally, vulnerability refers to the "*the quality or state of being exposed to the possibility of being attacked or harmed, either physically or emotionally*." (Lexico). It may also mean a weakness that can be exploited by one or more threats or a pre-disposition to suffer damage; or, it can be understood as the "*diminished capacity of an individual or group to anticipate, cope with, resist and recover from the impact*" (International Federation of Red Cross and Red Crescent Societies). Vulnerability varies with time (i.e., characteristics, driving forces, levels) Vogel & O'Brien (2004); DFID (2004). It is the anti-thesis of 'resilience' - which is the ability of an individual, a household, a community, a country or a region to withstand, to adapt, and to quickly recover from stresses and shocks (European Commission 2012).

There are various categorisations of vulnerable groups (in scholarship and policy). One of the more extensive ones is the *EquiFrame* conceptualisation of vulnerable groups which has 12 categories (Mannan et al 2012): 1. Limited resources (referring to poor people or people living in poverty), 2. Increased relative risk for morbidity (referring to people with one of the top 10 illnesses, identified by WHO, as occurring within the relevant country), 3. Mother child mortality (referring to factors affecting maternal and child health (0–5 years)), 4. Women headed household (referring to households headed by a woman), 5. Children (with special needs), referring to children marginalized by special contexts, such as orphans or street children 6. Aged (referring to older age). 7. Youth (referring to younger age without identifying gender). 8. Ethnic minorities (referring to non-majority groups in terms of culture, race or ethnic identity), 9. Displaced populations (referring to people who, because of civil unrest or unsustainable livelihoods, have been displaced from their previous residence), 10. Living away from services (referring to people living far from health services, either in time or distance), 11. Suffering from chronic illness (referring to people who have an illness which requires continuing need for care) 12. Disabled (referring to persons with disabilities, including physical, sensory, intellectual or mental health conditions, and including synonyms of disability).

More specifically, according to Andorno (2016), "*In human rights discourse for instance, the term vulnerability is used to indicate a heightened susceptibility of certain individuals or groups to being harmed or wronged by others or by the state. Populations which are particularly prone to being harmed, exploited or discriminated include, among others, children, women, older people, people with disabilities, and members of ethnic or religious minority groups.*" Andorno further elaborates, "*This does not mean that these groups are being elevated above others. Characterizing them as 'vulnerable' simply reflects the hard reality that these groups are more likely to encounter discrimination or other human rights violations than others*" – this is very relevant to our discussion as all of these categories are implicated in some form or manner in the legal issues and human rights principles at stake.

The use and deployment of AI technologies disproportionately affects vulnerable groups. E.g., The UNESCO COMEST *Preliminary Study on the Ethics Of Artificial Intelligence* gives an example of the Allegheny Family Screening Tool (AFST), a predictive model used to forecast child neglect and abuse. It states that it "*exacerbates existing structural discrimination against the poor and has a disproportionately adverse impact on vulnerable communities*" via oversampling of the poor and using proxies to understand and predict child abuse in a way that inherently disadvantages poor working families. Beduschi 2020 raises concerns about "*increasingly relying on technology to collect personal data of vulnerable people such as migrants and refugees*," to "*create additional bureaucratic processes that could lead to exclusion from protection.*" There are other examples. Children are particularly vulnerable (Butterfield-Firth 2018). As, the ICO explains, "*they may be less able to understand how their data is being used, anticipate how this might affect them, and protect themselves against any unwanted consequences*" (ICO undated). Individuals from the LGBTIQ[6] community might find themselves adversely affected by systems that permit or facilitate such profiling or discrimination. AI-powered data-driven and intensive economies might be more lucrative or attractive targets for cyberattacks given their expansive use of, and dependence on AI and big data.

In the AI context, vulnerability depends on various factors such as:

- **Physical/Technical,** e.g., poor design and/or development of algorithms and/or AI systems; inadequate security/protection; safety measures;
- **Social,** e.g., (lack of) public information and awareness about AI and its impacts, measures to ensure/protect well-being of individuals, communities and society, literacy, education, skills training, existence of peace and security, access to basic human rights, social equity, positive values, health, disabilities, social cohesion.
- **Political,** e.g., limited policy recognition of/strategy to address AI risks, preparedness measures, systems of good governance, incentives, e.g., to promote use of risk mitigation measures
- **Regulatory,** e.g., legislation, monitoring, enforcement, effective remedies for harms
- **Economic,** e.g., resources to cope with adverse effects, prosperity/poverty, investments in safe and ethically compliant systems, income levels, insurance.

The following table illustratively maps the identified AI legal issues to vulnerable groups and highlights the factors that determine and/or facilitate vulnerability

Tables 1 and 2

The above vulnerable groups are recognised to varying degrees in policy and regulatory discussions, but it can be argued that not enough is being done to protect them vis a vis taking effective action to prevent harms by addressing the factors of vulnerability themselves. Even where this is being done and there are some good steps being taken (e.g., at the EU-level and national level), it is far from where we need to be. So, how can the identified vulnerable groups be protected? Three actions are most required:

1 **Reduce** the adverse impacts of AI where possible through (continuous) risk identification, prediction, and preparation in consultation with affected stakeholders including a good representation of identified as vulnerable. This should be done at early stages in the research, design and development of AI technologies and evaluation of such measures
2 **Develop** and build capacities of vulnerable communities for resilience to such effects, and
3 **Tackle** the root causes of the vulnerabilities itself, e.g., taking a harder policy and regulatory stance on the harms, discrimination, inequality and injustice fueled by such technologies.

Action 1 is addressed directly to all actors in the AI ecosystem (researchers, research funders, developers, deployers, users, policy-makers). Action 2 is addressed to public policy-makers (at international, EU and national levels. Action 3 is addressed at regulators (all levels). Of the three actions, actions 2 and 3 are of immediate and urgent priority (since developments show we are addressing Action 1 to some extent, though this depends on context, applications and jurisdictions).

Conclusion

This article provided a panoramic overview of the myriad legal issues, gaps and challenges and affected human rights principles that are connected to AI and will function as particularly useful reference and stepping-stone for researchers conducting further studies on the topic

[6] Lesbian, gay, bisexual, transgender/transsexual, intersex and queer/questioning.

Table 1
Issues and affected human rights

AI legal issue	Human rights principles that might be affected
Lack of algorithmic transparency	fair trial and due process; effective remedies; social rights and access to public services; rights to free elections
Cybersecurity vulnerabilities	the right to privacy; freedom of expression and the free flow of information
Unfairness, bias and discrimination	elimination of all forms of discrimination against women; equal rights of men and women; enjoyment of children's rights without discrimination; equality before the law, equal protection of the law without discrimination; enjoyment of prescribed rights without discrimination; non-discrimination, right to life of migrant workers; right to liberty and security of the person; prohibition of discrimination on the basis of disability; right to fair trial; right to freedom from discrimination
Lack of contestability	right to an effective remedy; access to justice
Legal personhood, subjecthood, moral agency	right to recognition everywhere as a person before the law; right to equality; elimination of all forms of discrimination
Intellectual property issues	right to own property alone or in association with others; right to freely to participate in the cultural life of the community, to enjoy the arts and to share in scientific advancement and its benefits; right to the protection of the moral and material interests resulting from any scientific, literary or artistic production of which s/he is the author.
Adverse effects on workers	right to social security; prohibition of discrimination in relation to the enjoyment of rights to work, to free choice of employment, to just and favourable conditions of work, to protection against unemployment, to equal pay for equal work, to just and favourable remuneration; right to work, including the right of everyone to the opportunity to gain his living by work which s/he freely chooses or accepts); right of persons with disabilities to work, on an equal basis with others
Privacy and data protection issues	migrant's right to privacy; respect for privacy of person with disabilities; right to respect for private and family life; right to privacy and data protection; children's privacy; protection of the integrity of older persons and their privacy and intimacy
Liability issues related to damage caused	right to life; right to effective remedies
Lack of accountability for harms	right to life; right to effective remedies

Table 2
Mapping issues to vulnerabilities

Legal issue	Examples of most vulnerable group	Factors that determine/facilitate vulnerability (examples)
Lack of algorithmic transparency	People denied jobs, refused loans, refused entry/deported, imprisoned, put on no-fly lists or denied benefits.	Poor/bad/rogue design, unfit models. Ineffective regulation.
Cybersecurity vulnerabilities	SMEs/individuals with increased/ increasing reliance and dependence on AI-enabled technology. People in AI-powered data-driven and intensive economies. Children and youth.	Poorly designed and secured tech. Lack of resources. Investment and dependence on AI and data-driven technologies.
Unfairness, bias and discrimination	Ethnic/racially/gender stereotyped/profiled groups and minorities. Poor/low-income earners. Students allocated low grades and denied entry to educational opportunities.	Creator bias. Lack of consideration of ethical issue/focus on ethical design/lack of outputs testing and validation. Lack of provisions for human intervention.
Lack of contestability	Data subjects who lack the information they need to exercise rights.	Lack of information needed to exercise rights.
Legal personhood, subjecthood, moral agency	Humans whose rights and freedoms are affected/might conflict or compete.	Ill-considered policy and attribution of personhood.
Intellectual property issues	Inventors, creators of AI works.	Lack of clarity in provisions.
Adverse effects on workers	Young workers. Freelance/self-employed workers.	Lack of re-skilling and re-training. Inadaptable/inflexible education system.
Privacy and data protection issues	Children, disabled and/or older persons.	Dependence on AI and data-driven technologies.
Liability issues related to damage caused	Users of AI systems/those subject to AI use/persons to whom harm is caused e.g., in health/medical – disabled, chronically ill.	Overdependence on AI-powered technologies.
Lack of accountability for harms	Users of AI systems/those subject to AI use/persons to whom harm is caused especially civilians harmed in international AI-powered attacks.	Culture of non-accountability – lack of expectations and use of such standards. Use of exceptions/exemptions to bypass use of accountability promoting measures (above and/or within the law). No lasting consequences.

– in particular it connected the discussion of AI legal issues with vulnerability – a discussion that is much needed at many levels. Further, it presented three key actions that should be considered to protect vulnerable members of society.

Many of the examined issues have wide-ranging societal and human rights implications. They affect a spectrum of human rights principles: data protection, equality, freedoms, human autonomy and self-determination of the individual, human dignity, human safety, informed consent, integrity, justice and equity, non-discrimination, privacy and self-determination. The results of a socio-economic impact assessment carried out in the SIENNA project also highlighted concerns about such issues Jansen (2018). In addition to the specific issue-related challenges covered in this article, there are some general legal challenges – e.g., few AI specific regulations, lack of new regulatory bodies where existing ones fall short, sufficiency of existing national laws, lack of clarification on the application of existing laws, lack of legal academic debates in

some countries, lack of judicial knowledge and training, greyness in the legal status of automated systems Rodrigues (2019).

As AI technologies works closely with vast amounts of data, they will have cross-over and multiplicative effects that exacerbate legal and human rights issues related to them and impacts on individuals Rodrigues (2019). Such issues will amplify if industry develops applications and systems without paying attention early-on in the design and development process to the potential impacts of such technologies – whether on human rights, ethical and societal values (i.e., no use is made of privacy or ethics by design, ELSI analysis, impact assessments[7]).

With regard to the gaps, three themes repeat: A policy and legal shortfall, a technical shortfall and a multi-stakeholder shortfall in relation to AI. The *policy and legal* shortfall are being addressed to some extent (especially at the EU-level – see Rodrigues 2019), but at the same time caution and vigilance is required. The *technical* shortfall needs more serious consideration as it is at the point of technology design and development that the best positive influencing and requirements embedding can be done to address legal and ethical issues - well-designed AI would be half the battle won. The *multi-stakeholder shortfall* is tricky with different stakeholders bringing their own motivations to the table that need to be clearly understood (some to innovate unrestrictedly, others to ensure ethical and legal compliance, others to reap the profits of innovation in AI). Further the vulnerable and underrepresented community voices are not being heard enough. Still, a multi-stakeholder approach is being underlined (see e.g., Miailhe (2018) and addressed particularly at the international and EU levels.

Groups and communities most affected by such issues will vary depending on the context, application and use of AI, as shown in section 6. There is a critical need to tackle the factors that cause vulnerability head-on: by reducing the adverse impacts, developing capacities for resilience and tackling the root causes of the vulnerabilities.

As AI technologies progress, there will be further (and even amplified) legal issues, vulnerabilities and impacts on human rights that will need further monitoring and research. Technological advances will charge ahead via data-driven innovation and intelligent machines that complement and/or supplant the human and human capabilities. AI is at the forefront of discussions at the moment, but we expect the convergence of the technologies (AI, robotics, IoT) and new developments will change this, and refreshed discussions will be needed as new unique dilemmas for the law and our societal values will be posed.

Funding

This article draws from and builds upon the legal analysis results of the SIENNA project (Stakeholder-informed ethics for new technologies with high socio-economic and human rights impact) - which has received funding under the European Union's Horizon 2020 research and innovation programme under grant agreement No 741716.

Disclaimer

This article and its contents reflect only the views of the authors and does not intend to reflect those of the European Commission. The European Commission is not responsible for any use that may be made of the information it contains.

Acknowledgements

The author would like to thank the reviewers who provided feedback during article review and during the SIENNA research underpinning this article.

[7] E.g., privacy/data protection impact assessments, human rights impact assessments, algorithmic impact assessments.

Copyright

References

Access Now (2018) Human rights in the age of artificial intelligence. https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf.

Almada, M (2019). Human Intervention in Automated Decision-Making: Toward the Construction of Contestable Systems. In *17th International Conference on Artificial Intelligence and Law (ICAIL 2019)*. https://doi.org/10.2139/ssrn.3264189.

Aloisi, A, & Gramano, E (2019). Artificial intelligence is watching you at work. digital surveillance, employee monitoring and regulatory issues in the EU context. in: Stefano VD (ed) automation, artificial intelligence and labour protection. *Special Issue of Comparative Labor Law & Policy Journal*.

Ananny, M, & Crawford, K (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society, 20*(3), 973–989. https://doi.org/10.1177/1461444816676645.

Andorno, R (2016). Is vulnerability the foundation of human rights? *Human dignity of the vulnerable in the age of rights* (pp. 257–272). Cham: Springer.

Bartlett, M (2019). Solving the AI accountability gap. Hold developers responsible for their creations. *Medium*. https://towardsdatascience.com/solving-the-ai-accountability-gap-dd35698249fe.

Bathaee, Y (2018). The artificial intelligence black box and the failure of intent and causation. *Harvard Journal of Law & Tech, 31*(2), 889–937.

Bayamlıoğlu, E (2018). Contesting automated decisions. *European Data Protection Law Review, 4*, 433–446.

Beduschi, A (2020). International migration management in the age of artificial intelligence. *Migration Studies*, mnaa003. https://doi.org/10.1093/migration/mnaa003.

Berendt, B, & Preibusch, S (2017). Toward accountable discrimination-aware data mining: The importance of keeping the human in the loop—and under the looking glass. *Big data, 5*(2), 135–152.

Berk, RA (2019). Accuracy and fairness for juvenile justice risk assessments. *Journal of Empirical Legal Studies*. https://crim.sas.upenn.edu/sites/default/files/Berk_FairJuvy_1.2.2018.pdf.

Boden, MA (2016). *AI: Its Nature and Future*. UK: Oxford University Press.

Bodo, B, et al. (2018). Tackling the algorithmic control crisis-the technical, legal, and ethical challenges of research into algorithmic agents. *Yale Journal of Law and Tech, 19*(1), 3.

Brożek, B, & Jakubiec, M (2017). On the legal responsibility of autonomous machines. *Artificial Intelligence and Law, 25*(3), 293 30.

Brundage M (2018) The malicious use of artificial intelligence: forecasting, prevention, and mitigation. https://arxiv.org/pdf/1802.07228.pdf™sa=D™ust=1550739471109000.

Bryson, JJ, Diamantis, ME, & Grant, TD (2017). Of, for, and by the people: the legal lacuna of synthetic persons. *Artificial Intelligence and Law, 25*(3), 273–291.

Burri, T (2017). International law and artificial intelligence. *German Yearbook of International Law, 60*, 91–108. https://doi.org/10.2139/ssrn.3060191.

Butterfield-Firth, K (2018). Generation AI: What happens when your child's friend is an AI toy that talks back? *World Economic Forum*. https://www.weforum.org/agenda/2018/05/generation-ai-what-happens-when-your-childs-invisible-friend-is-an-ai-toy-that-talks-back/.

Cath, C (2018). Governing artificial intelligence: Ethical, legal and technical opportunities and challenges. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. https://royalsocietypublishing.org/doi/full/10.1098/rsta.2018.0080.

Čerka, P, Grigienė, J, & Sirbikytė, G (2017). Is it possible to grant legal personality to artificial intelligence software systems. *Computer Law & Security Review, 33*(5), 685–699.

CNIL (2017) How Can Humans keep the upper hand? The ethical matters raised by algorithms and artificial intelligence. Report on the public debate led by the French Data Protection Authority (CNIL) as part of the ethical discussion assignment set by the digital republic bill. https://www.cnil.fr/sites/default/files/atoms/files/cnil_rapport_ai_gb_web.pdf.

Coglianese, C, & Lehr, D (2018). Transparency and algorithmic governance. *Administrative Law Review*, 71.

Coldewey, D (2018). AI desperately needs regulation and public accountability, experts say. *Techcrunch*. https://techcrunch.com/2018/12/07/ai-desperately-needs-regulation-and-public-accountability-experts-say/?guccounter=1™guce_referrer_us=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8™guce_referrer_cs=YUKRqAi2qwbpTf7KOM4-jg. Accessed 11 August 2020.

Couchman, H (2019). Policing by machine. *Predictive Policing and the threats to our rights.*. https://www.libertyhumanrights.org.uk/sites/default/files/LIB%2011%20Predictive%20Policing%20Report%20WEB.pdf. Accessed 11 August 2020.

Council of Europe (2017) Recommendation 2102. Technological convergence, artificial intelligence and human rights. https://assembly.coe.int/nw/xml/XRef/Xref-XML2HTML-en.asp?fileid=23726™lang=en.

Council of Europe (2018) Discrimination, artificial intelligence, and algorithmic decision-making, Study by Prof. Frederik Zuiderveen Borgesius, Professor of Law, Institute for Computing and Information Sciences (iCIS), Radboud University Nijmegen, and Researcher at the Institute for Information Law, University of Amsterdam (the Netherlands. https://rm.coe.

int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73.

(2019). Declaration by the committee of ministers on the manipulative capabilities of algorithmic processes. In *Adopted by the Committee of Ministers on 13 February 2019 at the 1337th meeting of the Ministers' Deputies* https://search.coe.int/cm/pages/result_details.aspx?objectid=090000168092dd4b.

Courtland, R (2018). Bias detectives: the researchers striving to make algorithms fair. *Nature.*. https://www.nature.com/articles/d41586-018-05469-3.

Cummings, ML (2017). Artificial Intelligence and the future of warfare. *Research Paper.*. https://www.chathamhouse.org/sites/default/files/publications/research/2017-01-26-artificial-intelligence-future-warfare-cummings-final.pdf.

Danks, D, & London, AJ (2017). Algorithmic bias in autonomous systems. In *Proceedings of the 26th international joint conference on artificial intelligence*. AAAI Press https://www.cmu.edu/dietrich/philosophy/docs/london/IJCAI17-AlgorithmicBias-Distrib.pdf.

Dave, P (2018). Fearful of bias, Google blocks gender-based pronouns from new AI tool. *Reuters.*. https://www.reuters.com/article/us-alphabet-google-ai-gender/fearful-of-bias-google-blocks-gender-based-pronouns-from-new-ai-tool-idUSKCN1NW0EF.

Davies, CR (2011). An evolutionary step in intellectual property rights–Artificial intelligence and intellectual property. *Computer Law & Security Review, 27*(6), 601–619.

De Stefano, V (2019). Negotiating the algorithm': Automation, artificial intelligence and labour protection. *Comp. Labor Law & Policy Journal., 41*, 1.

Delcker, J (2018). Europe divided over robot 'personhood'. *Politico.*. https://www.politico.eu/article/europe-divided-over-robot-ai-artificial-intelligence-personhood/.

Desai, DR, & Kroll, JA (2017). Trust but verify: A guide to algorithms and the law. *Harv. JL & Tech, 31*, 1.

Department for International Development (DFID). (2004). Disaster risk reduction: a development concern. A Scoping Study on links between disaster risk reduction,. *Poverty and Development*. https://www.preventionweb.net/files/1070_drrscopingstudy.pdf.

Dignum, V (2018). The art of AI—accountability, responsibility, transparency. *Medium*. https://medium.com/@virginiadignum/the-art-of-ai-accountability-responsibility-transparency-48666ec92ea5.

Dizikes, P (2019). AI, the law, and our future. *MIT "Policy Congress" examines the complex terrain of artificial intelligence regulation*. http://news.mit.edu/2019/first-ai-policy-congress-0118.

Doshi-Velez, F, Kortz, M, Budish, R, Bavitz, C, Gershman, S, O'Brien, D, Schieber, S, Waldo, J, Weinberger, A, & Wood, A (2017). Accountability of AI under the law: The role of explanation. *arXiv preprint arXiv:1711.01134*, 2017. https://arxiv.org/pdf/1711.01134.pdf.

Edwards, L, & Veale, M (2017). Slave to the algorithm? Why a "right to an explanation" is probably not the remedy you are looking For. *Duke Law and Tech Review, 16*, 18–84.

European Commission (2012) Communication from the Commission to the European Parliament and the Council – The EU Approach to Resilience: Learning from Food Security Crises, COM (2012) 586 Final, Brussels, 3.10.2012. http://ec.europa.eu/echo/files/policies/resilience/com_2012_586_resilience_en.pdf.

Commission, European (2018a). Coordinated plan on artificial intelligence. In *Communication from the commission to the european parliament, the council, the european economic and social committee and the committee of the Regions, COM (2018) 795 final, Brussels, 7.12.2018.*

(2018). European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment. In *Adopted at the 31st plenary meeting of the CEPEJ* 3-4 December 2018 https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c.

(2018b). Communication from The Commission To The European Parliament. In *The European council, the council, the european economic and social committee and the committee of the regions. artificial intelligence for Europe.{SWD(2018) 137 final}. Brussels, 25.4.2018 COM(2018) 237 final* http://www.europarl.europa.eu/RegData/docs_autres_institutions/commission_europeenne/com/2018/0237/COM_COM(2018)0237_EN.pdf.

European Commission Expert Group on Liability and New Technologies – New Technologies Formation (2019). Liability for Artificial Intelligence and other emerging digital technologies. https://ec.europa.eu/transparency/regexpert/index.cfm?do=groupDetail.groupMeetingDoc™docid=36608.

(2020). *White Paper on artificial intelligence - a European approach to excellence and trust. Brussels, 19.2.2020 COM(2020) 65 final* https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf.

(2016). Artificial intelligence, robotics, privacy and data protection. In *Background document for the 38th international conference of data protection and privacy commissioners* https://edps.europa.eu/data-protection/our-work/publications/other-documents/artificial-intelligence-robotics-privacy-and_en.

European Data Protection Supervisor (EDPS) (2016) Artificial Intelligence, Robotics, Privacy and Data Protection, Background document for the 38th International Conference of Data Protection and Privacy Commissioners. https://edps.europa.eu/data-protection/our-work/publications/other-documents/artificial-intelligence-robotics-privacy-and_en.

European Data Protection Supervisor (EDPS) (2018) Opinion 5/2018 Preliminary opinion on Privacy by Design. https://edps.europa.eu/sites/edp/files/publication/18-05-31_preliminary_opinion_on_privacy_by_design_en_0.pdf.

European Parliament (2017) Resolution of 14 March 2017 on fundamental rights implications of big data: privacy, data protection, non-discrimination, security and law-enforcement (2016/2225(INI)).

European Parliament (2018a) Resolution of 12 February 2019 on a comprehensive European industrial policy on artificial intelligence and robotics (2018/2088(INI)).

European Parliament (2018b) Resolution of 25 October 2018 on the use of Facebook users' data by Cambridge Analytica and the impact on data protection (2018/2855(RSP)).

European Parliament (2019) Resolution of 15 January 2019 on autonomous driving in European transport (2018/2089(INI)).

European Parliament (2020a). Draft Report with recommendations to the Commission on a Civil liability regime for artificial intelligence (2020/2014(INL)). https://www.europarl.europa.eu/doceo/document/JURI-PR-650556_EN.html?redirect.

European Parliament (2020b). Draft Report with recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies (2020/2012(INL)). https://www.europarl.europa.eu/doceo/document/JURI-PR-650508_EN.html?redirect.

European Parliament (2020c). Draft Report on intellectual property rights for the development of artificial intelligence technologies (2020/2015(INI)). https://www.europarl.europa.eu/doceo/document/JURI-PR-650527_EN.html?redirect.

Feldstein, S (2019). The global expansion of AI surveillance. *Carnegie Endowment.*. https://carnegieendowment.org/2019/09/17/global-expansion-of-ai-surveillance-pub-79847.

Ferguson, D, & Savage, M (2020). Controversial exams algorithm to set 97% of GCSE results. *The Guardian*. https://www.theguardian.com/education/2020/aug/15/controversial-exams-algorithm-to-set-97-of-gcse-results. Accessed 15 August 2020.

Forbes Insights Team (2019) Rethinking Privacy For The AI Era. https://www.forbes.com/sites/insights-intelai/2019/03/27/rethinking-privacy-for-the-ai-era/.

Fralick, C (2019). Artificial intelligence in cybersecurity is vulnerable. *SC Magazine*. https://www.scmagazine.com/home/opinion/artifical-intelligence-in-cybersecurity-is-vulnerable/.

Frontier Economics (2018) The impact of artificial intelligence on work: an evidence review prepared for the royal society and the British academy. https://royalsociety.org/-/media/policy/projects/ai-and-work/frontier-review-the-impact-of-AI-on-work.pdf.

Gardner, S (2016). AI poses big privacy and data protection challenges. *Bloomberg Law News*. https://www.bna.com/artificial-intelligence-poses-n57982079158/.

Gluyas L, Day S (2018) Artificial Intelligence - Who Is Liable When AI Fails To Perform? CMS Cameron McKenna Nabarro Olswang LLP. https://cms.law/en/GBR/Publication/Artificial-Intelligence-Who-is-liable-when-AI-fails-to-perform.

Government of Canada (undated) Algorithmic Impact Assessment. https://canada-ca.github.io/digital-playbook-guide-numerique/views-vues/automated-decision-automatise/en/algorithmic-impact-assessment.html.

Guszca, J, Iyad, R, Bible, W, Cebrian, M, & Katval, V (2018). Why we need to audit algorithms. *Harv. Business Review.*. https://hbr.org/2018/11/why-we-need-to-audit-algorithms.

Hacker, P (2018). Teaching fairness to artificial intelligence: Existing and novel strategies against algorithmic discrimination under EU law. *Common Market Law Review, 55*(4), 1143–1185.

Hallevy G (2015) AI v. IP - Criminal Liability for Intellectual Property IP Offenses of Artificial Intelligence AI Entities. http://dx.doi.org/ 10.2139/ssrn.2691923.

High-Level Expert Group on Artificial Intelligence (AI HLEG) (26 June 2019) Policy and investment recommendations for trustworthy AI, https://ec.europa.eu/digital-single-market/en/news/policy-and-investment-recommendations-trustworthy-artificial-intelligence.

High-Level Expert Group on Artificial Intelligence (AI HLEG) (July 2020). The Assessment List For Trustworthy Artificial Intelligence (ALTAI) for self-assessment. https://ec.europa.eu/digital-single-market/en/news/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment Accessed 17 July 2020.

Hildebrandt, M (2016). The new imbroglio. *Living with machine algorithms* (pp. 55–60). The Art of Ethics in the Information Society.

House of Commons Science and Technology Committee (2018) Algorithms in decision-making. Fourth Report of Session 2017–19. https://publications.parliament.uk/pa/cm201719/cmselect/cmsctech/351/351.pdf.

IEEE Standards Association. P7003 - Algorithmic Bias Considerations. https://standards.ieee.org/project/7003.html.

IEEE Standards Association. P7001 - Transparency of Autonomous Systems. https://standards.ieee.org/project/7001.html.

Information Commissioner's Office (ICO) (2017) Big data, artificial intelligence, machine learning and data protection. Version: 2.2, 2017. https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf.

Information Commissioner's Office (ICO) (undated) When do we need to do a DPIA? https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/data-protection-impact-assessments-dpias/when-do-we-need-to-do-a-dpia/.

International Bar Association Global Employment Institute (2017) Artificial Intelligence and Robotics and Their Impact on the Workplace. https://www.ibanet.org/Document/Default.aspx?DocumentUid=c06aa1a3-d355-4866-beda-9a3a8779ba6e.

Jansen, P (2018). SIENNA. *D4.1 State of the art review: AI and robotics*. SIENNA project https://www.sienna-project.eu/digitalAssets/787/c_787382-l_1-k_sienna-d4.1-state-of-the-art-review–final-v.04-.pdf.

Jaynes T, L (2020). Legal personhood for artificial intelligence: citizenship as the exception to the rule. *AI & SOCIETY, 35*(2), 343 35.

Kingston, JKC (2016). Artificial intelligence and legal liability. In *International conference on innovative techniques and applications of artificial intelligence* (pp. 269–279). Cham: Springer.

Lepri, B, et al. (2018), *Fair, transparent, and accountable algorithmic decision-making processes* (31, pp. 611–627). Phil. & Tech..

Lindsey, N (2018). Artificial intelligence: privacy and legal issues. *CPO Magazine*. https://www.cpomagazine.com/data-privacy/artificial-intelligence-privacy-and-legal-issues/.

Liu, HW, Lin, CF, & Chen, YJ (2019). Beyond state v loomis: Artificial intelligence, government algorithmization and accountability. *International Journal of Law and Information Technology, 27*(2), 122–141.

Lohr, S (2019). AI and privacy concerns get white house to embrace global cooperation. *The New York Times*. https://www.nytimes.com/2019/04/03/technology/artificial-intelligence-privacy-oecd.html.
Mannan, Hasheem, et al. (2012). Core concepts of human rights and inclusion of vulnerable groups in the United Nations convention on the rights of persons with disabilities. *Alter, 6.3*(2012), 159–177.
Marr, B (2019). Artificial intelligence has a problem with bias, here's how to tackle it. *Forbes*. https://www.forbes.com/sites/bernardmarr/2019/01/29/3-steps-to-tackle-the-problem-of-bias-in-artificial-intelligence/.
Meyer, S (2018). Artificial intelligence and the privacy challenge. *CPO Magazine*. https://www.cpomagazine.com/data-privacy/artificial-intelligence-and-the-privacy-challenge/.
Miailhe, N (2018). AI & global governance: Why we need an intergovernmental panel for artificial intelligence. *AI & Global Governance*. https://cpr.unu.edu/ai-global-governance-why-we-need-an-intergovernmental-panel-for-artificial-intelligence.html.
Mitchell, I (2019). The use of AI gives rise to huge potential legal issues. *The Scotsman*. https://www.scotsman.com/lifestyle/iain-mitchell-the-use-of-ai-gives-rise-to-huge-potential-legal-issues-1-4924962.
Mittelstadt, BD, Allo, P, Taddeo, M, Wachter, S, & Floridi, L (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*. https://doi.org/10.1177/2053951716679679.
Niestadt, M, Debyser, A, Scordamaglia, D, & Pape, M (2019). Artificial intelligence in transport. In *Current and future developments, opportunities and challenges. Briefing paper* European Parliament Research Service http://www.europarl.europa.eu/RegData/etudes/BRIE/2019/635609/EPRS_BRI(2019)635609_EN.pdf.
Niiler, E (2019). Can AI be a fair judge in court? Estonia thinks so. *Wired*. https://www.wired.com/story/can-ai-be-fair-judge-court-estonia-thinks-so/.
OECD. (2019). *Artificial intelligence in society*. Paris: OECD Publishing.. https://doi.org/10.1787/eedfee77-en.
OECD (2019) Recommendation of the Council on Artificial Intelligence. Adopted on 22/05/2019. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.
Osoba, OA, & Welser IV, W (2017). The risks of artificial intelligence to security and the future of work. *RAND Corporation Santa Monica*. https://www.rand.org/pubs/perspectives/PE237.html.
Pasquale, F (2015). *The black box society, the secret algorithms that control money and information*. Harvard University Press.
Patel, A, Hatzakis, T, Macnish, K, Ryan, M, & Kirichenko, A (2019). D1.3 Cyberthreats and countermeasures. *SHERPA Project.*. https://doi.org/10.21253/DMU.7951292.v3.
Price, WN, II (2017). Artificial intelligence in health care: applications and legal implications. *The SciTech Lawyer, 14.1*. https://repository.law.umich.edu/cgi/viewcontent.cgi?article=2932™context=articles.
Privacy International and Article 19 (2018) Privacy and Freedom of Expression In the Age of Artificial Intelligence. https://www.article19.org/wp-content/uploads/2018/04/Privacy-and-Freedom-of-Expression-In-the-Age-of-Artificial-Intelligence-1.pdf.
Rachum-Twaig, O (2020). Whose robot is it anyway? Liability for artificial-intelligence-based robots. *University of Illinois Law Review, 2020*. Forthcoming. SSRN https://ssrn.com/abstract=3339230.
Raji, ID, & Buolamwini, J (2019). Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial AI products. *AAAI/ACM Conf. on AI Ethics and Society.*.
Raymond, AH, & Shackelford, SJ (2013). Technology, ethics, and access to justice: should an algorithm be deciding your case. *Mich. J. Int'l L., 35*, 485.
Reisman D, Schultz J, Crawford K, Whittaker M (2018) Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability. https://ainowinstitute.org/aiareport2018.pdf.
Ribeiro, MT, Singh, S, & Guestrin, C (2016). Why should I trust you?: Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1135–1144). ACM.
Rigby, MJ (2019). Ethical dimensions of using artificial intelligence in health care. *AMA Journal of Ethics, 21*(2), 121–124.
Rodrigues R (2019) SIENNA D4.2: Analysis of the legal and human rights requirements for AI and robotics in and outside the EU, SIENNA project. https://www.sienna-project.eu/digitalAssets/801/c_801912-l_1-k_sienna-d4.2-legal-analysis-ai-robotics-awaiting-approval.pdf.
Roig, A (2017). Safeguards for the right not to be subject to a decision based solely on automated processing (Article 22 GDPR). *European Journal of Law and Tech, 8*, 3.
Schönberger, D (2018). Deep Copyright: Up-and downstream questions related to artificial intelligence (AI) and machine learning (ML). *Zeitschrift fuer Geistiges Eigentum/Intellectual Property Journal, 10*(1), 35 5.
Siemaszko, K, Rodrigues, R, & Slokenberga, S (2020). D5.6: Recommendations for the enhancement of the existing legal frameworks for genomics, human enhancement, and AI and robotics. *SIENNA project*.
Smith, L (2017). Unfairness by algorithm: Distilling the harms of automated decision-making. *Future of Privacy Forum.*. https://fpf.org/2017/12/11/unfairness-by-algorithm-distilling-the-harms-of-automated-decision-making/.
Talking Tech. (2017). AI and IP: copyright in AI-generated works (UK law). *Can copyright subsist in an AI-generated work?*. https://talkingtech.cliffordchance.com/en/ip/copyright/ai-and-ip–copyright-in-ai-generated-works–uk-law-.html.
The UK Copyright Service (2004) Fact sheet P-15: Designs and design rights. https://www.copyrightservice.co.uk/protect/p15_design_rights.
Tschider, C (2018). Regulating the IoT: Discrimination, privacy, and cybersecurity in the artificial intelligence age. *Denv. U. L. Rev., 96*, 87.
Turner, J (2019). Legal personality for AI. *Robot Rules* (pp. 173–205). Cham: Palgrave Macmillan.
United Nations (2019) United Nations Activities on Artificial Intelligence (AI). http://handle.itu.int/11.1002/pub/813bb49e-en Accessed 11 August 2020.
Vaccaro K, Karahalios K (undated) Algorithmic Appeals. https://s3.amazonaws.com/kvaccaro.com/documents/algappeal.pdf.
Vayena, E, Blasimme, A, & Cohen, IG (2018). Machine learning in medicine: Addressing ethical challenges. *PLoS medicine, 15*(11), Article e1002689.
Vladeck, DC (2014). Machines without principals: liability rules and artificial intelligence. *Wash. L. Rev., 89*, 117.
Vogel, C, & O'Brien, K (2004). *Vulnerability and global environmental change: rhetoric and reality*. AVISO 13. In *Global environmental change and human security project*.
Wachter, S, Mittelstadt, B, & Floridi, L (2017). Transparent, explainable, and accountable AI for robotics". *Science Robotics, 2*(6), eaan6080.
Wachter, S, & Mittelstadt, BD (2019). A right to reasonable inferences: Re-thinking data protection law in the age of Big Data and AI. *Columbia Business Law Review*. https://ora.ox.ac.uk/objects/uuid:d53f7b6a-981c-4f87-91bc-743067d10167/download_file?file_format=pdf™safe_filename=Wachter%2Band%2BMittelstadt%2B2018%2B-%2BA%2Bright%2Bto%2Breasonable%2Binferences%2B-%2BVersion%2B6%2Bssrn%2Bversion.pdf™type_of_work=Journal+article.
Wallace, N (2017). EU's right to explanation: A harmful restriction on artificial intelligence. *Tech Zone, 360*. https://www.techzone360.com/topics/techzone/articles/2017/01/25/429101-eus-right-explanation-harmful-restriction-artificial-intelligence.htm.
Williams, H (2019). Big brother AI is watching you. *IT ProPortal*. https://www.itproportal.com/features/big-brother-ai-is-watching-you/.
World Intellectual Property Organization (WIPO) (1998) Intellectual Property and Human Rights", proceedings of a panel discussion, organized by the World Intellectual Property Organization in collaboration with the Office of the United Nations High Commissioner for Human Rights, on 9 November 1998. http://www.wipo.int/edocs/pubdocs/en/intproperty/762/wipo_pub_762.pdf.

Further reading

Lexico (2020) Vulnerability. https://www.lexico.com/definition/vulnerability.
McKinsey Analytics (2019) Global AI Survey: AI proves its worth, but few scale impact. https://www.mckinsey.com/i/media/McKinsey/Featured%20Insights/Artificial%20Intelligence/Global%20AI%20Survey%20AI%20proves%20its%20worth%20but%20few%20scale%20impact/Global-AI-Survey-AI-proves-its-worth-but-few-scale-impact.pdf Accessed 11 August 2020.
PwC (2027) Sizing the prize. What's the real value of AI for your business and how can you capitalise?. https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf Accessed 11 August 2020.
SHERPA project. https://www.project-sherpa.eu/about/ Accessed 11 August 2020.
UNESCO COMEST (2019) *Preliminary Study on the Ethics Of Artificial Intelligence*, SHS/COMEST/EXTWG-ETHICS-AI/2019/1 Paris, 26 February 2019. https://unesdoc.unesco.org/ark:/48223/pf0000367823.
CEIPI (2019) "Artificial Intelligence and intellectual property". http://www.ceipi.edu/en/training-seminars/artificial-intelligence-and-intellectual-property/ Accessed 11 August 2020.
Centre for Information Policy Leadership (2018) Artificial Intelligence and data protection: delivering sustainable AI accountability in practice first report: Artificial intelligence and data protection in tension. https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_ai_first_report_-_artificial_intelligence_and_data_protection_in_te....pdf Accessed 11 August 2020.
(March 2018). Study on the human rights dimensions of automated data processing techniques (in particular algorithms) and possible regulatory implications. In *Prepared by the committee of experts on internet intermediaries (MSI-NET)* https://rm.coe.int/study-hr-dimension-of-automated-data-processing-incl-algorithms/168075b94a.
Council of Europe signatories. (2018). Justice by algorithm – the role of artificial intelligence in policing and criminal justice systems. *Motion for a recommendation*. Doc. 14628. 26 September 2018 https://www.coe.int/en/web/commissioner/-/safeguarding-human-rights-in-the-era-of-artificial-intelligence.
Edwards, L, & Veale, M (2018). Enslaving the algorithm: from a 'right to an explanation' to a 'right to better decisions'? *IEEE Security and Privacy Magazine, 16*(3), 46–54. https://doi.org/10.1109/MSP.2018.2701152.
EU Agency for Fundamental Rights (FRA) (2018) #BigData: Discrimination in data-supported decision making. http://fra.europa.eu/en/publication/2018/big-data-discrimination.
European Commission (2018) Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions on Artificial Intelligence for Europe. https://ec.europa.eu/digital-single-market/en/news/communication-artificial-intelligence-europe.
(2020). *Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics, COM(2020) 64, February 2020.*
(2016). Artificial Intelligence, robotics, privacy and data protection. In *Background document for the 38th international conference of data protection and privacy commissioners* https://edps.europa.eu/data-protection/our-work/publications/other-documents/artificial-intelligence-robotics-privacy-and_en.
Baeza-Yates, R (2016). Data and algorithmic bias in the web. In *Proceedings of the 8th ACM Conference on Web Science*.
European Parliament and the Council, Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data,

and repealing Directive 95/46/EC (General Data Protection Regulation), *OJ L* 119, 4.5.2016, p. 1–88.

European Parliament (2017) Resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL).

European Parliament (2018c), Resolution of 12 September 2018 on autonomous weapon systems (2018/2752(RSP)).

European Parliament, Resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), Official Journal of the European Union, C 252/239, 18.7.2018.

European Parliamentary Research Service (2019) A governance framework for algorithmic accountability and transparency. http://www.europarl.europa.eu/RegData/etudes/STUD/2019/624262/EPRS_STU(2019)624262_EN.pdf.

Hajian, S, Bonchi, F, & Castillo, C (2016). Algorithmic bias: From discrimination discovery to fairness-aware data mining. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining.* ACM.

Hallevy, G (2010). The criminal liability of artificial intelligence entities - from science fiction to legal social control. *Akron Intellectual Property Journal, 4*, 2. Article 1 http://ideaexchange.uakron.edu/akronintellectualproperty/vol4/iss2/1.

House of Lords Select Committee on Artificial Intelligence (2018) Report of Session 2017–19. AI in the UK: ready, willing and able? https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf.

International Federation of Red Cross and Red Crescent Societies (2020) What is vulnerability? https://www.ifrc.org/en/what-we-do/disaster-management/about-disasters/what-is-a-disaster/what-is-vulnerability/.

Committee on Civil Liberties, Justice and Home Affairs, Report on fundamental rights implications of big data: privacy, data protection, non-discrimination, security and law-enforcement (2016/2225(INI). http://www.europarl.europa.eu/doceo/document/A-8-2017-0044_EN.html?redirect.

European Data Protection Supervisor (EDPS). Artificial Intelligence. https://edps.europa.eu/data-protection/our-work/subjects/artificial-intelligence_en.

IBM Research Trusted AI. AI Fairness 360 Open Source Toolkit. https://aif360.mybluemix.net.

European Commission (undated) Liability for defective products. https://ec.europa.eu/growth/single-market/goods/free-movement-sectors/liability-defective-products_en.