Barriers to Regulating AI: Critical Observations from a Fractured Field

Ashlin Lee[1] (ORCID iD: 0000-0003-0524-121X), Will Orr[3] (ORCID iD: 0000-0002-6720-5794), Walter G. Johnson[2] (0000-0001-6359-2312), Jenna Imad Harb[2], and Kathryn Henne[2] (ORCID iD: 0000-0001-7369-3002)

[1] Environmental Informatics Group, Land and Water, CSIRO, Canberra, Australia; School of Sociology and Justice and Technoscience Lab, The Australian National University, Canberra, Australia
[2] Justice and Technoscience Lab, School of Regulation and Global Governance (RegNet), The Australian National University, Canberra, Australia
[3] Annenberg School of Communication, University of Southern California, Los Angeles, United States of America; Justice and Technoscience Lab, The Australian National University, Canberra, Australia.

Abstract
This chapter reflects upon the current state of artificial intelligence (AI) regulation, observing how both AI and its regulation are fragmented and fractured. We observe this fracturing across three axes. The first axis concerns definitional fractures, drawing attention to the fragmented state of AI as a concept and as an object of regulation. The second axis considers regulatory interventions, particularly how regulatory responses have not been unified or consistent in responding to the development and spread of AI. The final axis points to the unequal dispersal of benefits and burdens, and how AI regulation unevenly treats different actors. Drawing on Jackson's concept of repair work, we argue that these fractures, while being evidence of sub-optimal regulation, also provide new opportunities to reflect on the underlying dynamics of AI and its regulation in context.

Keywords:
Artificial Intelligence (AI); race to AI regulation; fracture; responsive regulation; repair; self-regulation

Introduction

Experts increasingly call for the regulation of artificial intelligence (AI) on the grounds that it can create risks and harms that disproportionally impact marginalised groups (Buolamwini & Gebru, 2018; Raji et al., 2020). The competitive desire to create more powerful AI—sometimes called the "AI arms race"—facilitates these harms by encouraging the development of unvetted AI systems (Scharre, 2019). Growing evidence supports concerns about AI-related harms. For example, the use of AI to sort and judge individuals in the context of employment opportunities or the provision of welfare benefits can perpetuate bias and discrimination (Eubanks, 2018). In response, a "race to AI regulation" (Smuha, 2021) has emerged, with government and corporate actors developing responses, including frameworks, guidelines, policies, and standards, to address a growing range of AI-related concerns (Floridi & Cowls, 2019).

This chapter reflects critically on the state of AI regulation. Regulation broadly refers to steering the flow of events and is undertaken by various public and private actors including, but not limited to, government (Parker & Braithwaite, 2003). Here, we highlight how the current state of AI regulation is shaped by a significant degree of fracturing and fragmentation. We observe this fracturing across three axes. The first axis concerns definitional fractures, drawing attention to the fragmented state of AI as a concept and as an object of regulation. The second axis considers regulatory interventions, particularly how regulatory responses have not been unified or consistent in responding to the development and spread of AI. The final axis points to the uneven dispersal of benefits and burdens, and how AI regulation unevenly treats different actors.

Drawing on insights from critical infrastructure studies, critical data studies, and regulatory governance, we outline the implications of fragmentation and fracturing in the context of AI regulation. Instead of critiquing the state of AI as evidence of failure, we see these breakages as valuable leverage points. In some, though not all, instances, there is value in maintaining fractures to pursue more just technological outcomes. We highlight how fractures can add value by looking at the challenges of AI regulation through the lens of breakage and repair, which Jackson (2014) describes as the work required for systems to operate over time through moments of disruption.

Fractured Definitions: AI as a Fragmented Concept

Fragmented definitions regarding how AI is used, and what problems AI can address, complicate the development of regulatory approaches. At a technical level, the computational systems and mathematics required for AI are complex, with many different approaches and uses. Efforts to respond to AI are also diverse, with a wide range of regulatory actors and methods responding to the risks of AI. For example, government, non-government, and private sector organisations have all made suggestions for regulating AI (Jobin et al., 2019), proposing approaches ranging from industry-led self-regulation (Taeihagh, 2021) to legally enforceable bans and moratoriums (The United Nations, 2021). Fragmented understandings of both what AI is and how it might be regulated complicates opportunities for effective regulation.

Conceptually, what constitutes AI is quite broad. For example, Nilsson (2009) describes AI as about making machine intelligence, with intelligence describing foresight and action in one’s environment. General definitions thus tend to focus on computer systems possessing intelligent capacities – such as learning and pattern recognition – that are usually found in humans. A focus on intellectual capacities, however, is somewhat distorting, as there are fundamental differences in how humans and machines form and execute these capacities. Such a definition reveals little about these differences and what AI systems actually look like and do (Fjelland, 2020). To better clarify AI, Corea (2019) suggests AI might be defined by considering:

- What problems is AI used to solve, including reasoning, knowledge, planning, communication, and perceptual challenges;
- The computational approach used to address these problems;[1] and,

[1] Corea (2019) further breaks down this category into detailed subcategories describing the specific mathematical and computational techniques used. Examples given include: logic-based tools using programmed rules; knowledge-based tools using ontological databases of rules and relationships; probabilistic tools using incomplete information to make decisions through statistics; machine learning which uses pattern recognition and reinforcement approaches; embodied intelligence that connects sensory perception and

- Whether AI is narrow and focused on a specific problem, or general and capable of multiple uses.

General references to AI and common synonyms, such as "machine learning," "neural networks," and "deep learning," all refer to specific technical characteristics and usage circumstances. However, they are often used without clarification, further fuelling its fragmentation. Definitions of AI are therefore messy, heterogenous, and unclear.

Without deeper investigation, important elements of AI's definition and regulation can be missed, causing further fragmentation. For example, the algorithmic models that underlie AI tend to be considered the most valuable aspect of AI (Sambasivan, 2022). However, such a view obscures crucial aspects of AI, such as the work required to create and curate high-quality datasets for AI models. Without such work, AI cannot exist, and ignoring this work can result in the failure of AI techniques and applications. This work can also be highly problematic, with data workers frequently devalued and exploited in the data supply chain (Gray & Suri, 2019). For regulation to be extensive it is thus important to understand the key mathematical and algorithmic components of AI, *and* the broader set of supporting elements (such as the production of training datasets) that allow AI to function.

Because AI has a broad set of current and future applications, incremental approaches that suggest qualified principles and specific guidelines have preferred over more general regulatory paradigms (Reed, 2018). In the absence of longitudinal data and evidence, scholars have argued these principles and guidelines are largely normative in nature, targeting the general modus operandi of AI vis-à-vis human values rather than specific risks that AI might present (Wirtz et al., 2020). The significant growth in "Ethical AI" guidelines offers one such example. Created by both public and private sector actors, Ethical AI guidelines attempt to set ethical values and expectations for AI's development and use, in the hope that systemic risks and potential social, economic, and political harms might be avoided (Boddington, 2017). These values and standards, however, are not necessarily legally binding or enforceable.

Surveying the landscape, Jobin and colleagues (2019) found 84 examples of national AI ethics guidelines, nearly equally divided amongst private and public organisations but with notable underrepresentation from Central Asia, Africa, and Southern and Central America. These frameworks converge around five themes: transparency, justice and fairness, non-maleficence, responsibility, and privacy. Floridi and Cowls (2019) similarly isolated five principles in their analysis of AI ethics standards adopted by major public and private actors: beneficence, non-maleficence, autonomy, justice, and explicability. de Almeida and colleagues (2021) identified 21 regulatory approaches that adopted variations of these principles in their design. They note that, despite convergence on high-level principles of Ethical AI, significant variations remain.

Despite greater availability of ethical AI principles, examples of successful regulation using these principles remain scarce. The technical complexity of AI means any regulation must be

---

interactions with intelligent activities such as virtual reality; and, finally, search and optimisation that attempts to optimise responses to search queries. Although current at the time of publication, it is also conceivable that new techniques or variations have evolved given the rapid pace of development.

flexible enough to accommodate the diversity of AI, while also acknowledging the diversity of contexts in which AI is applied. Ethical AI guidelines attempt to do this by focusing on normative principles and abstract concepts of ethics. In this way, Ethical AI guidelines may be useful in forming the basis of stronger regulatory responses, as they capture a "normative core" of values for AI (Fjeld et al., 2020). However, despite the abundance of Ethical AI guidelines and principles, it remains to be seen whether these guidelines can be enforced and if they provide sufficient oversight to avoid AI-related harms.

Fractured Interventions: Limitations in Regulatory Interventions

Regulatory interventions toward AI are also fragmented. Despite the diversity of guidelines developed around AI, a lack of clear advice on implementation remains. While fragmentation in regulatory interventions is common (Black, 2002), definitional uncertainty around AI compounds this problem, as the process of creating and applying new norms and standards requires some common ground. Fragmentation in regulatory interventions can occur along multiple dimensions, including: (1) the type of institutional setting implementing norms, (2) level of government, (3) legal bindingness, (4) regulatory target, and (5) broad or specific construction of a norm (see Abbott & Snidal, 2009). While a comprehensive review of these interventions is beyond the scope of this chapter, a brief overview here illustrates their fragmentation and its implications.

The unrestrained growth of regulatory actors and measures has fragmented the capacity of regulation to influence AI, particularly through the predominance of self-regulation as the preferred approach to regulating AI. Using Ethical AI as the primary lens for regulating AI, technology firms (e.g., Microsoft, Meta) have introduced their own ethical codes. Multistakeholder bodies (e.g., World Economic Forum [WEF], Organisation for Economic Cooperation and Development [OECD]), and public entities including legislatures, regulators, and all levels of courts, have also contributed different guidelines. Non-state actors as diverse as civil society bodies, professional organizations, industry groups, and standard-setting organisations have also begun regulatory initiatives. These include codes of conduct, technical standards, and certification or monitoring programs. For example, the Institute of Electrical and Electronics Engineers CertifAIEd Program (IEEE, 2022) provides a risk-based framework and ethics criteria to certify the development and operation of AI systems as compliant on issues including transparency, bias, accountability, and privacy. The practical capacity of these different sectoral and jurisdictional approaches is contingent on the assumption that AI designers and operators will actively and responsibly self-regulate. This assumption is, however, challenged in two ways. First, through evidence that Big Tech players actively oppose most regulation, lobbying against it on the basis that it impedes innovation (Satariano & Stevis-Gridneff, 2020), and second that the guidelines that underpin the technology industry's preferred approach of self-regulation (i.e., AI ethics guidelines) are non-binding and unenforceable. Big Tech actors materially benefit from fragmentation, as they can position self-regulation as legitimate and practical, and oppose more stringent and enforceable government regulations.

Through promoting self-regulation, corporate actors come to have unique obligations and capacities to craft and enforce regulation. Compliance with many of these codes of conduct remains legally voluntary and open to (subjective) interpretation. In addition, such codes often lack monitoring and accountability provisions (Bowman & Hodge, 2009). Instead, compliance and enforcement of the ethical principles rely more heavily on reputational and market forces, as well as internal normative motives. The presence of various ethical codes may further fragment regulatory interventions by presenting an array of choices and by decreasing the perceived need for states to issue binding regulation. In doing so, they can contribute to different states around the world codifying rules of varying strength at different points in time. For instance, state-based regulators have more legally binding options available than non-state or hybrid entities, though government institutions have often been reticent to set binding rules on AI to date. One such example is how the United States relies on a largely voluntary regulatory program for AI in autonomous vehicles (McAslan et al., 2021). This program sets non-binding standards and reporting norms for developers to manage the application of AI in autonomous vehicles. In contrast, the European Union appears poised to enact a more comprehensive and binding regulatory framework for AI, adopting a risk-based approach for AI across sectors.

While many non-state interventions are not legally binding, indirect pathways to enforcement may still be available through liability, insurance, or consumer protection regulation—adding further fragmentation (Wallach & Marchant, 2019). These interventions can also lack the checks and balances expected of state-based interventions. For example, where state-based interventions often require procedural justice, freedom of information requests, and public record-keeping, non-state actors are not necessarily subject to these requirements. Seeking accountability is thus not always clear or straightforward.

Definitional uncertainty fuels further fragmentation in the practical implementation of AI regulation. This regulatory environment yields unclear obligations, limited options for accountability, and varied perceptions of AI regulation amongst both key stakeholders and the public. This is concerning, as it can erode the belief in enforceable AI regulation.

Fractured Outcomes: The Uneven Distribution of Burdens and Benefits of AI Regulation

The fractured landscape of AI regulation means that its benefits and burdens are unevenly distributed. This relates to a reliance on Ethical AI guidance over enforceable regulation, and the “AI arms race” logic that is driving the growth of AI. For example, Radu (2021) argues that the regulation of AI has suffered through the “hybridisation of governance”, where external stakeholders (including think tanks, corporations, and other actors) are enlisted to shape the direction of AI governance and regulation. Wishing to support the growth of AI, nation-states defer regulatory responsibility to external stakeholder groups – hybridising governance between the interests of stakeholders and the public.

Without an authoritative position on regulation set by the government, however, these groups are allowed to define the formal and informal rules of regulation for the state with little direction or mandate. For instance, national AI strategies, designed with stakeholder input, often focus on high-level Ethical AI guidance and self-regulation rather than enforceable regulatory measures (Floridi & Cowls, 2019). This creates a disorderly regulatory environment that cements power amongst those already invested in AI (such as commercial entities), while making it difficult for those outside these privileged groups to contribute their knowledge and experience. For instance, as Sloane (2022) notes, the European Commission's AI Alliance, an online forum for providing feedback on AI policy decisions in the EU, quickly devolved into an unrepresentative echo-chamber without capturing the lived experiences of those impacted, or even the broader view of industry and unaligned experts. As Radu (2021, p. 190) argues "it becomes increasingly hard to disentangle public interest policies from market dominance interests." External stakeholders stand to benefit from the hybridised regulation of AI, with the public potentially less well served by this arrangement.

The popularity of self-regulation in the private sector is another example of the hybridisation of governance. Framing the regulation of AI through loose ethical principles has encouraged corporate actors to adopt and promote self-regulation as an approach that limits the risks of AI while not impeding its potential innovations (Greene et al., 2019). In many tech companies, new roles have been established to manage the ethical impacts of AI systems (Metcalf et al., 2019). These roles often attempt to regulate AI systems internally within existing corporate protocols, such as review boards and codes of conduct. Self-regulation allows corporations to appear compliant with loose regulations by defining when and how Ethical AI guidelines can be applied while keeping costs to a minimum (Metcalf et al., 2019).

The atmosphere of self-regulation has provided opportunities for ethics-washing. Ethics-based rhetoric and institutionally acceptable interventions are mobilised to frame business-as-usual practices as ethical, while avoiding substantive action that might disrupt the interests of the corporation (Wagner, 2019). Industry-led self-regulations like Ethical AI can therefore be used to create a competitive advantage, to build credibility and increase trust, thereby increasing revenue (Wagner, 2019). In short, "[p]rofit maximisation … is rebranded as bias minimisation" (Benjamin, 2019, p. 30). As such, the self-regulation of AI allows for ethical concerns to be captured by corporate logics of "meritocracy, technological solutionism, and market fundamentalism" (Metcalf et al., 2019, p. 470), and managed through reductive and performative ethics washing measures that lack real-world effectiveness and any real ethical value (Bietti, 2021).

The broad and often ambiguous nature of many AI regulations can mean negotiating disagreements often fall to the judicial system. For instance, a school in Sweden was fined in violation of the European Union's General Data Protection Regulation (GDPR) for unlawfully testing facial recognition technology. As per the GDPR, testing systems do not usually require approval by data regulators, as subjects' consent is considered a sufficient threshold for the processing of biometric data (Penner & Chiusi, 2020). In this case, however, student consent was not deemed to be freely given due to the power disparities between the institution and its pupils. However, as Galanter (1974) argues, the rules and structures of the judicial system often reinforce existing power imbalances between those who can afford lengthy court processes and access to legal expertise, and those who cannot. Unlike the general public and those who suffer the consequences of AI, Big Tech actors and AI developers can afford multiple court cases, fines, and legal fees incurred while fighting regulation, making regulation a cost of business, rather than an effective deterrent. While essential, the judicial system may not be the best place to negotiate the regulation of AI given it can advantage those with the most resources and least relative risk and while disadvantaging those with the least resources and greatest relative risk.

Frontline operators, engineers, data scientists, and other technical practitioners often manage the burden to interpret and enact regulatory principles. Practitioners are required to translate ethical and regulatory principles into tangible technical constraints in the design of AI systems (Orr & Davis, 2020). Although practitioners can use tools to aid the translation of ethical and regulatory codes into design protocols, many codes are imprecise in nature. Designers prefer more detailed, technical standards set by standard-setting bodies, which allow professional certification of their work and pathways towards industry recognised standards, such as the European Union's CE mark, which indicates conformity with EU safety and quality requirements (e.g., Henriksen et al., 2021). Thus, even when tools are available to practitioners, they can hinder designers' ability to adapt the restrictions to the realities of developing AI (Morley et al., 2020). Loopholes also potentially allow practitioners to continue to accumulate data, further benefiting their position, while the consequences of loopholes fall disproportionately on consumers, civil society, or those most at risk. Practitioners often have an assumed responsibility for the ethical integrity of their system, and a wide-ranging capacity to ensure this integrity. However, ethical responsibility for determining how, by whom, and to what effect AI is regulated is distributed throughout the sociotechnical system of actors (Orr & Davis, 2020). This fractured landscape muddies clear lines of accountability as to "who makes the rules and for how long" (Radu, 2021, p. 190).

Fragmentation also shapes how the benefits of AI regulation are distributed. Technical hurdles – such as limitations in how transparent and easily understood an AI's foundational algorithms are (Amoore, 2020) – can undermine efficiency and reach of regulatory interventions. This causes the potential benefits of regulation – such as preventing malfeasance, ensuring accountability, and other normative principles described in section 1 – to become dispersed amongst a limited number of groups, rather than universally applied. As discussed in section 2, the coverage and quality of interventions vary, fracturing the supposed benefits of regulation. For instance, while the GDPR has enshrined the right for data subjects to receive "meaningful information about the logic involved" (The European Union, 2022a) in making automated decisions, this standard is difficult to achieve. It should serve the interest of anyone subject to automated decisions, while protecting those who are most likely to experience algorithmic harms. Yet, whether it is technically possible for these benefits to materialise is unclear. Even the most explainable AI requires expert knowledge to understand (Burrell, 2016), and technical transparency is not sufficient for understanding a system's logic or for ensuring accountability (Amoore, 2020). Feminist analyses underscore that technical solutions often occlude the systemic and structural factors underlying algorithmic discrimination and calls for fairness (West, 2020). Furthermore, systems motivated by vague notions of 'social good' are at risk of reinforcing logics of colonial extraction and domination (Madianou, 2021), failing to address the power imbalances stemming from AI (Gebru, 2020).

Beyond the technical limitations of these regulations, caveats in legislation undermine their effectiveness. For instance, regulation on automated decision-making in the GDPR only applies to "decisions based solely on automated processing" (The European Union, 2022b) such as profiling. If a decision is made by both AI and a human, the "logic" behind these decisions need not be provided to data subjects. For example, an Uber driver's accusations of algorithmically-enabled unfair dismissal were deemed illegitimate in a Dutch court due to a human agent making the final decision (Lomas, 2021). This example raises important questions about how much power AI's have over human beings as part of collaborative decision systems. Furthermore, regulatory limitations on algorithmic power only apply to decisions that produce legal or similarly significant effects for the individual, such as the automated refusal of credit applications. Non-legal harms, such as racial bias through AI and the loss or limitation of disability access, are less easily captured. Structural power imbalances in the development and implementation of regulation thus drive the unequitable distribution of benefits relating to AI and its regulation (Wachter et al., 2017).

Discussion

This chapter has drawn attention to challenges presented by the fractured global landscape of AI regulation. Reflecting on these fractures in the context of regulatory theory, it is natural to assume that promoting the mending of these fractures is the logical next step for regulation. For example, Braithwaite (2017), when explaining responsive regulation, argues for a shared basis of learning and capacity building between regulatory actors as the base of what he calls the "regulatory pyramid." Responsive regulation dynamically responds to the needs of government, private sector, and third-party actors through appropriate movements up or down the regulatory pyramid, which includes responses and sanctions. Before the enforcement of punitive measures, actors should be allowed to learn and build their capability to act appropriately. Shared bodies of knowledge on AI are, however, contested and not authoritative. Regulatory interventions are shared amongst stakeholders, who are independently moving up or down their own regulatory pyramids, as the common ground shared between actors (e.g., AI ethics guidelines) holds few binding mechanisms. This leads to the burden of regulation falling on individuals (designers and engineers, for example) or systems that were never intended to arbitrate society-wide issues of rapid technical change (such as the judiciary or corporate-driven self-regulation). Given the shortcomings of the current regulatory environment, it is therefore understandable that a desire to unify AI regulation and mend any fractures might be a productive and logical step forward.

Despite this logical connection between the fractured state of AI and a lack of successful regulation, we suggest that desires to unify AI regulation by plastering over these fractures are misplaced. Instead, a key to understanding and regulating AI is to embrace these fractures as windows into the details that underpin the field of AI. Through understanding these fractures and the associated work – what Jackson (2014) describes as "repair work" – we gain insights into how regulation might better respond to complex systems like AI. Jackson (2014) argues that greater attention should be paid to the breakdowns, failures, and disruptions of sociotechnical systems, and the work required to make them function at sites of failure. Breakages and fractures reveal the limits of how AI thinks, works, and operates. They are therefore sites where we might begin to change AI.

Using the lens of repair, fracturing can be approached as part of the solution for AI regulation. While vested interests might prefer the dual arms races of AI innovation and regulation to be administrated by experts and prestigious institutions, fractures reveal the messy, partial, and continuing work associated with the development of AI *and* the development of AI regulation. As Jackson (2014) argues, innovation is often framed as something that happens in a laboratory, where technology is perfected, and failure occurs well after innovation, separate from this process. This framing, however, does not reflect the reality of AI regulation. There is ongoing intellectual and discursive work to create guidance for the sector, which is sometimes contested. Deployment of guidelines and regulation means working with diverse, and sometimes disagreeable users, professionals, and policy makers. These moments of contestations reveal the social and technical elements most valuable to AI, and therefore those elements that regulation should attempt to engage.

A focus on fractures, fragmentation, and breakages may also help progress the regulation of AI by illuminating alternative views of AI that foreground issues of justice and equity. For example, Costanza-Chock's (2020) summary of the Design Justice movement emphasises end-to-end, participatory design methods with the communities that technologies like AI directly impact. This requires an acceptance of the fragmentation of AI and its regulation through different contexts and experiences that are not always accounted for in the design process. By embracing the partial, contextual, and sometimes contested accounts that different groups experience, the basis for a more nuanced technical and regulatory response to AI may be developed. Given the complexity of AI, desires to create unified paradigms of AI ethics or regulation give little room to address these nuances and the needs of those living with and through AI systems. As section 3 discusses, AI's burdens and benefits are often distributed unevenly amongst different groups, each with their own needs and ethical relationships, which are constantly shifting. A "one-size-fits-all" approach to AI regulation fails to respect these differences and their implications. In some cases, it may be undesirable or unadvisable to have uniform approaches to AI regulation, such as in instances where AI harms disproportionately affect certain demographic groups. In these instances, uniform and cohesive rules would not speak to the specific AI harms and risks these groups face, and would not achieve just outcomes. In this way, embracing fragmentation may be more productive.

This does not mean we reject some of the advancements achieved thus far. For instance, as noted earlier, there is some moderate convergence on the normative core of Ethical AI, and around key principles by which AI ethics should be structured. This normative core may form the basis of appropriate and powerful regulation in the future, with appropriate oversight and enforcement mechanisms. This normative core should not be treated, however, as consensus or a finished product, as ethical relationships are still very much at play through a lens of fracture. According to Jackson (2014), repair foregrounds forgotten ethical relationships of mutual care and responsibility between both humans and technology, and the chains of interactivities that create ethical relationships – as fulfilled through repair work. With AI systems exercising power over more aspects of everyday life, questions of what one owes, and is owed, as a part of their life become important. Exploring points of fracture provides opportunities to interrogate different norms, understand their real-world impact amongst those at risk, and craft alternatives better aligned with those in need.

Conclusion

AI and its regulation are—and will likely continue to be—fractured by virtue of competition amongst actors, a political-economic drive for innovation, diverse and shifting contexts in which AI is applied, and the varying technical complexities of AI. Embracing these fractures is a starting point for building more effective regulation. The fractured view of AI foregrounds the reality of AI development and deployment, and respects the diversity of challenges that different communities face. Regulating AI is not and will not be a one-size-fits-all endeavour, thus harmonised and standardised approaches to regulatory intervention are unrealistic. Instead, it is a constantly iterative process that must wrestle with competing views, social and technical limitations, and continuing power imbalances. Regulating AI is a difficult task, but it will be even more challenging if these fractures are ignored.

References

Abbott, K. W., & Snidal, D. (2009). The Governance Triangle: Regulatory Standards Institutions and the Shadow of the State. In *CHAPTER TWO. The Governance Triangle: Regulatory Standards Institutions and the Shadow of the State* (pp. 44–88). Princeton University Press. https://doi.org/10.1515/9781400830732.44

Amoore, L. (2020). *Cloud ethics: Algorithms and the attributes of ourselves and others*. Duke University Press.

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim code*. Polity.

Bietti, E. (2021). From Ethics Washing to Ethics Bashing: A Moral Philosophy View on Tech Ethics. *Journal of Social Computing*, *2*(3), 266–283. https://doi.org/10.23919/JSC.2021.0031

Black, J. (2002). Critical reflections on regulation. *Australian Journal of Legal Philosophy*, *27*, 1–35.

Boddington, P. (2017). *Towards a code of ethics for artificial intelligence*. Springer.

Bowman, D. M., & Hodge, G. A. (2009). Counting on codes: An examination of transnational codes as a regulatory governance mechanism for nanotechnologies. *Regulation & Governance*, *3*(2), 145–164. https://doi.org/10.1111/j.1748-5991.2009.01046.x

Braithwaite, J. (2017). Types of responsiveness. In *Regulatory Theory,: Foundations and Applications* (pp. 117–132). ANU Press.

Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of Machine Learning Research*, *81*, 1–15.

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, *3*(1), 205395171562251. https://doi.org/10.1177/2053951715622512

Corea, F. (2019). AI Knowledge Map: How to Classify AI Technologies. In F. Corea (Ed.), *An Introduction to Data: Everything You Need to Know About AI, Big Data and Data Science* (pp. 25–29). Springer International Publishing. https://doi.org/10.1007/978-3-030-04468-8_4

Costanza-Chock, S. (2020). *Design justice: Community-led practices to build the worlds we need*. MIT Press.

de Almeida, P. G. R., dos Santos, C. D., & Farias, J. S. (2021). Artificial Intelligence Regulation: A framework for governance. *Ethics and Information Technology*, *23*(3), 505–525. https://doi.org/10.1007/s10676-021-09593-z

Eubanks, V. (2018). *Automating inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin's Press.

Fjeld, J., Achten, N., Hilligoss, H., Nagy, A., & Srikumar, M. (2020). Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3518482

Fjelland, R. (2020). Why general artificial intelligence will not be realized. *Humanities and Social Sciences Communications*, *7*(1), 10. https://doi.org/10.1057/s41599-020-0494-4

Floridi, L., & Cowls, J. (2019). A Unified Framework of Five Principles for AI in Society. *Harvard Data Science Review*. https://doi.org/10.1162/99608f92.8cd550d1

Galanter, M. (1974). Why the "Haves" Come Out Ahead: Speculations on the Limits of Legal Change. *Law and Society Review*, *9*(1), 95–160.

Gebru, T. (2020). Race and Gender. In *The Oxford Handgook of Ethics of AI* (pp. 253–270). Oxford University Press.

Gray, M., & Suri, S. (2019). *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*. Houghton Mifflin Harcourt.

Greene, D., Hoffmann, A. L., & Stark, L. (2019). *Better, Nicer, Clearer, Fairer: A Critical Assessment of the Movement for Ethical Artificial Intelligence and Machine Learning*. https://doi.org/10.24251/HICSS.2019.258

Henriksen, A., Enni, S., & Bechmann, A. (2021). Situated Accountability: Ethical Principles, Certification Standards, and Explanation Methods in Applied AI. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 574–585). Association for Computing Machinery. https://doi.org/10.1145/3461702.3462564

IEEE. (2022). *IEEE CertifAIEd*. https://engagestandards.ieee.org/ieeecertifaied.html

Jackson, S. J. (2014). Rethinking Repair. In T. Gillespie, P. J. Boczkowski, & K. A. Foot (Eds.), *Media Technologies* (pp. 221–240). The MIT Press. https://doi.org/10.7551/mitpress/9780262525374.003.0011

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Lomas, N. (2021). Uber hit with default 'robo-firing' ruling after another EU labor rights GDPR challenge. *TechCrunch*. https://social.techcrunch.com/2021/04/14/uber-hit-with-default-robo-firing-ruling-after-another-eu-labor-rights-gdpr-challenge/

Madianou, M. (2021). Nonhuman humanitarianism: When "AI for good" can be harmful. *Information, Communication & Society*, *24*(6), 850–868. https://doi.org/10.1080/1369118X.2021.1909100

McAslan, D., Gabriele, M., & Miller, T. R. (2021). Planning and Policy Directions for Autonomous Vehicles in Metropolitan Planning Organizations (MPOs) in the United States. *Journal of Urban Technology*, *28*(3–4), 175–201. https://doi.org/10.1080/10630732.2021.1944751

Metcalf, J., Moss, E., & boyd, danah. (2019). Owning ethics: Corporate logics, silicon valley, and the institutionalization of ethics. *Social Research: An International Quarterly*, *86*(2), 449–476.

Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2020). From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices. *Science and Engineering Ethics*, *26*(4), 2141–2168. https://doi.org/10.1007/s11948-019-00165-5

Nilsson, N. J. (2009). *The Quest for Artificial Intelligence: A history of Ideas and Achievements*. Cambridge University Press.

Orr, W., & Davis, J. L. (2020). Attributions of ethical responsibility by Artificial Intelligence practitioners. *Information, Communication & Society*, *23*(5), 719–735. https://doi.org/10.1080/1369118X.2020.1713842

Parker, C., & Braithwaite, J. (2003). Regulation. In P. Cane & M. Tushnet (Eds.), *The Oxford Handbook of Legal Studies* (pp. 119–145). Oxford University Press.

Penner, K., & Chiusi, F. (2020). *Automating Society Report 2020—European Union*. Algorithm Watch. https://automatingsociety.algorithmwatch.org/report2020/european-union/

Radu, R. (2021). Steering the governance of artificial intelligence: National strategies in perspective. *Policy and Society*, *40*(2), 178–193. https://doi.org/10.1080/14494035.2021.1929728

Raji, I. D., Gebru, T., Mitchell, M., Buolamwini, J., Lee, J., & Denton, E. (2020). Saving Face: Investigating the Ethical Concerns of Facial Recognition Auditing. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 145–151. https://doi.org/10.1145/3375627.3375820

Reed, C. (2018). How should we regulate artificial intelligence? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *376*(2128), 20170360. https://doi.org/10.1098/rsta.2017.0360

Sambasivan, N. (2022). All equation, no human: The myopia of AI models. *Interactions*, *29*(2), 78–80. https://doi.org/10.1145/3516515

Satariano, A., & Stevis-Gridneff, M. (2020, December 14). Big Tech Turns Its Lobbyists Loose on Europe, Alarming Regulators. *The New York Times*. https://www.nytimes.com/2020/12/14/technology/big-tech-lobbying-europe.html

Scharre, P. (2019). Killer apps: The real dangers of an AI arms race. *Foreign Aff.*, *98*, 135.

Sloane, M. (2022). To make AI fair, here’s what we must learn to do. *Nature*, *605*(9), 1. https://doi.org/10.1038/d41586-022-01202-3

Smuha, N. A. (2021). From a ‘race to AI’ to a ‘race to AI regulation’: Regulatory competition for artificial intelligence. *Law, Innovation and Technology*, *13*(1), 57–84. https://doi.org/10.1080/17579961.2021.1898300

Taeihagh, A. (2021). Governance of artificial intelligence. *Policy and Society*, *40*(2), 137–157. https://doi.org/10.1080/14494035.2021.1928377

The European Union. (2022a). Art. 13 GDPR – Information to be provided where personal data are collected from the data subject. *General Data Protection Regulation (GDPR)*. https://gdpr-info.eu/art-13-gdpr/

The European Union. (2022b). Art. 22 GDPR – Automated individual decision-making, including profiling. *General Data Protection Regulation (GDPR)*. https://gdpr-info.eu/art-22-gdpr/

The United Nations. (2021, September 15). *Urgent action needed over artificial intelligence risks to human rights*. UN News. https://news.un.org/en/story/2021/09/1099972

Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the general data protection regulation. *International Data Privacy Law*, *7*(2), 76–99.

Wagner, B. (2019). Ethics As An Escape From Regulation. From "Ethics-Washing" To Ethics-Shopping? In E. Bayamlioglu, I. Baraliuc, L. A. W. Janssens, & M. Hildebrandt (Eds.), *Being Profiled: COGITAS ERGO SUM* (pp. 84–89). Amsterdam University Press. https://doi.org/10.1515/9789048550180-016

Wallach, W., & Marchant, G. (2019). Toward the Agile and Comprehensive International Governance of AI and Robotics [point of view]. *Proceedings of the IEEE*, *107*(3), 505–508. https://doi.org/10.1109/JPROC.2019.2899422

West, S. M. (2020). Redistribution and Rekognition: A Feminist Critique of Algorithmic Fairness. *Catalyst: Feminism, Theory, Technoscience*, *6*(2). https://doi.org/10.28968/cftt.v6i2.33043

Wirtz, B. W., Weyerer, J. C., & Sturm, B. J. (2020). The Dark Sides of Artificial Intelligence: An Integrated AI Governance Framework for Public Administration. *International Journal of Public Administration*, *43*(9), 818–829. https://doi.org/10.1080/01900692.2020.1749851