



Analyzing the Contribution of Ethical Charters to Building the Future of Artificial Intelligence Governance

Lyse Langlois¹(✉) and Catherine Régis²

¹ Faculty of Social Sciences, Laval University, Québec, Canada
Lyse.langlois@observatoire-ia.ulaval.ca

² Faculty of Law, University of Montreal, Montréal, Canada
catherine.regis@umontreal.ca

Keywords: Ethics · Ethical charters · Artificial intelligence · Formalisation of ethics · Social regulation · Norms · Legal framework

A technology is not merely a system of machines with certain functions; rather it is an expression of a social world (Nye 2007).

1 Introduction¹

Advances in digital technology and artificial intelligence (AI) systems have the potential to radically transform economies and societies. The deployment of these technologies promises to improve a host of important services for society as a whole, but also raises concerns with regard to uncontrolled development of technological systems and a certain inappropriate use or disregard for social factors. Technology is non-neutral (Gautrais 2012); it encompasses values and can even affect them. Technology has such power that it structures and defines the end-purposes of human activity, to the extent that individuals and communities have no choice but to adapt and change under its influence. AI systems (AIS) are also a vehicle for power struggles: they can help people build their capacities as well as constrain these by segmenting society and reinforcing social inequities and injustices (Barabas et al. 2018). In view of these potential digressions and the significant impact on societies, this fifth technological revolution has quickly led to widespread discussion (i.e., academic, but also political and very public) of the ethical issues raised by the exercise of (ir)responsible AI. Incorporating various AIS without ethical consideration of their impact² has already led to greater discrimination of certain groups, criticism likely to be biased, and

¹ The authors would like to express their gratitude to the external reviewers as well as professors Miriam Cohen and Bryn William Jones for their very useful comments on this chapter.

² As reference: COMPAS (Correctional Offender Management Profiling for Alternative Sanctions), a criminal recidivism prediction tool (2016), the Facebook job offer recommendation algorithm favouring male applicants over women (2019), the AppleCard, which discriminated against women applying for lines of credit (2019).

developers' and decision-makers' potential lack of ethical sensitivity, whether intentional or not (Bairaktarova and Woodcock 2017). As set forth in the Villani report, AIS cannot be allowed to be new instruments to exclude or overly track people without public debate (Villani 2018) and considerations for human rights. In view of these concerns, a certain number of organizations have developed ethical charters to govern the development of AI and related digital technologies. More strategic and ambitious reports have also been deployed to help shape or define the AI and digital technology policy of a state or a group of states.

Given the significant academic and public policy attention being given to the development of ethical charters to identify principles and values likely to provide a better framework for AI and digital technology, we argue that it is important to also more fully address the formalization of these charters. Formalization involves assessing when and how the principles and values in a charter can be institutionalized in organizations or professional practices. Formalization actually contributes to establishing AI ethics. And such an assessment provides fertile ground for experimentation on the use of AI ethics as a potential preamble to future digital rights or, on a larger scale, AI legal frameworks.

In the first part of this article, we examine the elements that appear essential to charter formalization based on an organizational ethics research approach (Murphy 1989; Adelman 1991; Carroll and Bucholtz 1999; Mercier 2004), in order to make a case for *ethics* as one component of developing AI governance. The second part of this article follows with the outcomes of our reflections in this regard. The legal framework of AI is increasingly being considered as the next major step in normative development in this sector, both at the local and international levels. We will explore potential synergies between ethical charter assets and legal developments likely to be deployed in building this next step.

2 The Ethical Charter Landscape: The First Component of AI Governance Development

Before exploring the formalization of ethical charters, there is a need to better understand what has been accomplished so far in the ethical charter landscape. While there has been valuable and useful progress in defining shared values and tools to guide AIS developments, we also saw some conduct that limits their impact and legitimacy. Acknowledging such complex ethical charter landscape, we then examine the elements needed to formalize ethics that encourage a beneficial materialization of such charters.

2.1 The Good, the Bad and the Ugly

The flourishing development of ethical AI principles throughout the world in the last few years has surely been constructive, but due to their multiplicity such statements of AI ethics principles have sometimes been confusing for the AI research community and the general public. The development of AI ethics principles has contributed to stimulating local and international discussions on AI benefits and risks from various cultural, professional and disciplinary perspectives. However, as we will point out later,

the voices of countries from the South and minority groups are still underrepresented in this movement. It has also elicited more formal positioning from various stakeholders (governments, industries, international organizations, academia, etc.) regarding what “good AI”³ means or does not mean, through such mechanisms as administrative directives, best practices, and organizational policies and regulations. The well-known work of Jobin et al. (2019) and Zeng (2019) implemented through a platform that tracks in real time every ethical charter initiative (<https://www.linking-ai-principles.org/>) has been useful in illustrating the points of convergence and divergence that emerge from these multiple ethical guidelines (e.g., 84 guidelines in 2020). Their work demonstrates that stakeholders’ views clearly converge on a few specifically shared principles, even if they do not necessarily agree on their definition. The work of Fjeld et al. (2020), from the Berkman Klein Center for Internet and Society, also demonstrates an effort to integrate and analyze charters from various regions in the world (Latin America, Middle East, etc.). They identified thirty-six relevant charters and ranked them using a typology that features various principles based on a proportion of convergence, their potential links with human rights and the groups of stakeholders concerned. For the authors, these points of convergence can represent the “normative core of a principle-based approach to AI ethics and governance” (2020:5).

Arguably, the fact that ethics is not binding has sparked more candid debates, at least in some circumstances, about AI challenges in various contexts. Such ethical guidelines were often viewed as necessary tools to encourage reflexive approaches, in order for data collectors, developers, researchers, users and policymakers to develop more responsible AI behaviour instead of using mandatory “checklists” likely to be applied uniformly but without further thought. This discussion could have been different, perhaps more cautious, if we had started by addressing legal reforms and considering their direct economic, political and social impact as well as their ensuing crystallization effect.

However, despite the fact that such ethical initiatives have galvanized AI ecosystem stakeholders since 2016, an opposite movement has emerged quite recently to undermine some of the initiatives. Referred to as “fake ethics” or ethics-washing/bashing, this phenomenon is defined by Elettra Bietti (2019) as a trend to discredit or trivialize AI ethics in general and governance initiatives in particular. There are different sources for such a phenomenon.

One of these has to do with the nature of the groups undertaking the development of AI ethics charters with interests sometimes in contradiction to the actual end-purposes of the initiatives. Several researchers (O’Neil 2016; Fontanel and Sushcheva 2019; Metzinger 2019) have revealed industry lobbies intent on delaying the implementation and development of AI regulations and policies. For example, some principles serving as recommendations in various discussion groups (NGOs, public authority and universities) were removed and replaced with less exacting principles that diluted agreements and resulted in generic, non-binding documents. The adoption of

³ The word *good* is in brackets as the very meaning of good AI is a complex issue: Ben Green, ““Good” isn’t good enough”, 2019, Neurips conference paper, available online: https://aiforsocialgood.github.io/neurips2019/accepted/track3/pdfs/67_aisg_neurips2019.pdf.

such principles thus orients the direction of debate and determines the type of innovation on which to focus in this sector while influencing rules and policies, especially when they serve as strategic documents. Moreover, the composition of groups of experts on the matter has been criticized as being politically driven and not always fairly representative of various groups (gender, sector, discipline, country, etc.). A certain discrepancy has become apparent among group representatives, with a minimal number of ethicists compared with private sector representatives. Attention to inequalities in representation have also highlighted the fact that a much greater number of individuals from high income countries are chosen in comparison to individuals from low or middle-income countries in the global south.⁴

Despite these considerations, it should be noted that some businesses have embraced ethics and even hired philosophers to set up committees or services often referred to as “AI and Ethics” or “AI for the Benefit of Humanity,” and for the most part these are led by legal experts. According to Metzinger (2019), most of these initiatives are part of communication strategies meant to uphold the reputations of organizations by promoting their good standing among the population. Hence, a certain ethical legitimacy is conveyed because ethicists were involved in the process. With regard to such initiatives—with some more akin to marketing strategies—the content of meaning covered by “ethics” can be considered of second- or even third-degree importance compared with other interests. Such tactics have often been criticized during processes deployed to institutionalize ethics (Salomon 2007).

The fact that ethical charters often leave a “practical gap” is another source of frustration. In other words, the ethical principles that they promote do not readily translate into technical solutions, particularly for developers (Floridi 2019). A study by Miller and Coldicott found that 79% of technology workers wished to have more practical resources to help them with ethical considerations (2019). As Morley et al. (2019) clearly put it, there is a current need for the AI ethics community to embark on a second, more practical endeavour, which requires translation from the “what” to the “how.” In line with the other critiques mentioned above, this gap is problematic. It can discredit valuable ethical progress and benchmarks, which can lead to their rejection by the technical and AI user community and to gradual disinterest in a future ethical conversation, sometimes referred to as “ethics shrinking” (Floridi 2019). In cases where people reflect, on an ongoing basis, about the benefits and risks of their practice developing at an extremely fast pace and for which it is almost impossible to predict every ensuing challenge, reflexive approaches meant to be developed through such charters could be lost.

The recent phenomenon of ethical initiatives aimed at providing a better framework for AI warrants sustained attention and its claims should be submitted to serious review. As the term “ethics” has long been used independent of academic discussion, over time it has gradually been divested of its content and meaning. Peters, Vold, Robinson and Calvo (2020), as well as Brent Mittelstadt (2019), agree on this concept:

⁴ See to this effect the awareness-raising paper produced by UNESCO, *Steering AI and advanced ICTs for knowledge societies: a Rights, Openness, Access, and Multi-stakeholder Perspective*, Xianhong Hu [1], Neupane, Bhanu, Echaiz, Lucia Flores, Sibal, Prateek, Rivera Lam, Macarena, 2019.

for charters to go beyond their declarative purposes, they must overcome the challenge of implementing the ethical principles that they encompass. As stated by Hagedorff (2020), “a very little to nothing has been written about the tangible implementation of ethical goals and values” (2020, p. 100). To avoid limiting ourselves to cataloguing good intentions, we deemed it necessary to highlight the formalization process of ethical charters and principles by considering them from the perspective of organizational ethics and social sciences and humanities research. Through this approach, the ethical and social imperatives found in these documents can be materialized and genuinely leveraged to transform professional practices, processes, organizational systems and legal frameworks.

2.2 The Formalization of Ethics: A Social Regulation Tool

The strong interest in organizational ethics is part of a movement driven by consumers, investors, wage earners and the population in general, with regard to an increasing social demand for a greater integration of ethics in organizational life (Mercier 1999). This movement started at the end of the 1980s in response to a crisis of trust regarding government institutions and the abuse of assets and public funds. The dynamics that led to these new perspectives involving applied ethics served to develop a body of knowledge at the origin of organizational ethics. From an instrumental perspective, organizational ethics focuses on how organizations integrate values into their policies, practices and decision-making processes. It also proposes a critical reflection on every aspect of an organization.

According to several organizational ethics researchers (Boisvert 2011; Bégin, 2009), the issue of trust is strongly related to the need to rely on ethics institutionalization. Under its commitment to fight corruption, the Organisation for Economic Cooperation and Development (OECD) also promoted organizational ethics at the end of the 1990s, by establishing a working group dedicated to the promotion of ethical infrastructure. It is also under the OECD’s impetus that organizational ethics research was developed and led to active collaborations between Quebec, French and Belgian researchers. These researchers have been particularly interested in the phenomenon of formalization and its impact on professionals and organizations. Formalization is more than just an aspect in the development of a more encompassing ethics institutionalization system within organizational governance (Mercier 1999:22). It can also translate the demands of civil society stakeholders for greater transparency in decision-making and better access to information. Moreover, the formalization of ethics reflects the willingness of organisations to go beyond the intention stage and implement in processes and practices an ethical framework to include a process of reflection and an action plan regarding strategies to mitigate risks in a technological design project. Ethical charters are often the first mechanisms to be put in place, because once formalised, they serve as guidelines and prescriptions in terms of rules of conduct. In addition, they can also have normative implications if they are clearly promoted in organizational processes and practices by senior management.

The phase of formalization raises a certain paradox often mentioned by ethicists and philosophers: institutionalizing ethics as a type of social regulation (Reynaud 1991) can reduce it to a prescriptive and normative dimension. This vision often sparks

controversy among philosophers: for them, ethics is above all a tool for critical reflection on values and end-purposes rather than a management tool. From the perspective of social regulation, ethics focus on the relationships established in practice between various axiological and normative registers with constraints, regulations, obligations and responsibilities. This vision proves entirely relevant when it is rooted in the field of applied ethics and when it allows to deliberate on values and principles. In this respect, decision-making has become the central issue meant to resolve situation clarification problems. It is in practice, at the core of AI design systems, that some issues and challenges come to light and cause both potential emancipation and exploitation to emerge. The implementation of risk mitigation strategies based on an ethical framework is part and parcel of a culture mindful of ethics and social responsibility. The social logic of applied ethics is essentially structured around the concept of responsibility. G. Legault identified that “ethical responsibility refers to the *response* given to those who question the *value* of decisions made in practice. (2000, p. 33) Hence, applied ethics can only be operated in concrete situations, in other words, on a case-per-case basis. By emphasizing the assessment of decisions *in situ*, applied ethics assert a local approach focussing on individuals whose impact on others and society.” (2000, p. 34) As a background, ethical questioning is not about knowing “how one should live”, a construct that refers to the question to which morality responds (morality dictates). Conversely, ethics respond to the question “how to live” (ethics recommend), which is the quite insightful argument of ethicist J. Dratwa (2019) featured in his book *Dans quel monde voulons-nous vivre ensemble?* All in all, the purpose of applied ethics is to find an answer to the challenges of AI development with a view to ensure and maintain the development of a technology for the common good.

For social sciences and humanities experts, however, ethics formalization is a form of acknowledgment of stakeholders’ aptitude to develop their ethical competency and ability for reflexivity, dialogue and autonomy: ethics can be a tool for empowerment (Langlois 2014). Studies conducted in this domain highlight the possibility for these dimensions to relate and interrelate, by formalizing ethics through its relationship with instruments like charters, on the one hand and its use, on the other hand, while preserving its reflexive nature. Hence, the formalization of ethics embodies the willingness of organizations to be socially responsible by promoting a form of ethical formalization related to AI development.

Dual-Purpose Ethical Institutionalization: An Object of Reflection and a Social Regulation Tool

The formalization of ethics is required in order to go beyond the stage of *good intentions* by giving an official, legitimate place to ethics within organizations and associations. Ethics can be integrated into various suitable internal mechanisms and modes of management, in order to ensure compliance with commitments stated or promoted by an organization as well as their external adherence by implementing relationships of trust with stakeholders.

For instance, an organization can adopt the principles of a technological development statement or ethical charter in order for such a commitment to positively influence strategic orientations and its planning, as well as to ensure an integrated involvement in the lifecycle of AIS. With regard to AI, the early days of ethics formalization focused

mainly on heightening awareness about risks and issues, followed by a will to integrate ethical principles as early as during algorithm design. Hence, injunctions were raised for this purpose (ethics by design, security by design, privacy by design). Integrating these principles prior to the design phase can raise new ethical obligations for designers and decision-makers, while also being indicative of a strong ethical commitment when these are actually considered upstream, implemented and promoted in AIS. Maintaining the dual purpose of ethics - a mode of social regulation and a reflexive capacity conducive to empowerment - can be seen as a dynamic process that contributes to consolidating the formalisation of ethics in AIS.

The Three Objectives of Ethics

It is the very nature of ethics to offer space for reflection while remaining relevant and flexible enough to respond to new challenges. Peters et al. (2020) mentioned that “[...] ethical impact evaluation must be an ongoing, iterative process—one that involves various stakeholders at every step, and can be re-evaluated over time, and as new issues emerge.” The philosopher Malherbe (2000) appropriately highlighted the two modes of cohabitation by stressing that ethics has three objectives: (1) to create spaces for deliberation to support professionals and the public in their reflection on values and norms; (2) to provide these people with tools to better measure the ethical outcomes of their decisions; and (3) to rethink work organization methods, professional practices, modes of management and societal life.

Malherbe’s proposal provides an interesting framework to analyze the AI charters deployed for ethical purposes. For instance, the first two objectives of ethics are addressed in The Montréal Declaration for a Responsible Development of Artificial Intelligence (2018), given that the Declaration provides access to the methodological process selected to validate its ethical principles. Based on an expert-citizen co-construction, this stage identifies the various phases of development built on a declaration of general ethical principles and fundamental values, such as well-being, autonomy, justice, privacy, knowledge, democracy and responsibility. Through an iterative phase focused on co-construction, several groups were invited to contribute to a reflection on those values while affirming the selected principles. At this time, there is insufficient information to measure the impact of this declaration on work organization methods, professional practices, modes of management and societal life. Given that the overall phenomenon of ethical charters is relatively recent,⁵ meeting this objective requires a formalization phase. However, the phase apparently the most akin to this third objective would possibly be Ethically Aligned Design (IEEE-2019⁶) which had its ethical principles converted into practice standards (IEEE-P7000-P7010). If these standards obtain some form of recognition through certification, they will have demonstrated their importance and will be able to affect modes of management and professional practices, such as those of engineers.

⁵ We rediscover the appearance of the first ethical charters in 2016 according to the typology of the Berkman Klein Center: Principled AI: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI, <https://nrs.harvard.edu/urn-3:HUL.InstRepos:42160420>.

⁶ IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems (n 5) pp. 21–22 (See Principle 2.).

The Need for Legitimacy in the Ethical Process

The formalization of principles and values contained in ethical charters challenges the impact of such documents on AI and digital frameworks. In turn, the legitimacy associated with those charters is also challenged. Based on Malouf's philosophical dictionary, legitimacy is defined as what "*complies with laws as well as morals and reason.*" [Translation] Malouf further specifies that "*legitimacy is what allows peoples and individuals to accept, without excessive constraint, the authority of an institution embodied by humans and considered as bearing shared values.*" [Translation] (Malouf 2009:2) The *Littre* dictionary defines this in terms of what is rooted in equity and reason. Habermas (1978) essentially addresses legitimacy by proposing a concept of communicational ethics. Our interest here is focused on the definition that preserves and maintains trust in AIS developments proposed by Rosavallon (2008), in which he addresses the importance of establishing a legitimacy of proximity. Based on dialogue, such legitimacy is built on the capacity of individuals or professionals⁷ to discuss and take part in such issues with a genuine impact on their lives. According to Bourgeois and Nizet (1995), "a behaviour, an opinion or a decision is legitimate for stakeholders if they perceive it as complying with social norms that they consider positive." [Translation] (1995:34) According to Luhmann (2006), such deliberation in the public sphere can "reduce social complexity" and, as a result, obtain stronger social acceptability. It is our view that, in this process to formalize ethical charters, the implementation of a legitimacy of proximity makes ethics a political choice. The common denominator of these charters is often the promotion of society's well-being and trust in technological developments. Therefore, it is all the more important that public deliberation be essential, given the end-purpose of these instruments. Therefore, it confirms its legitimacy through the compelling demonstration of its necessity.

Now that we have some insight of the ethical charter landscape and that we have explored the elements needed to formalize ethics that guarantee the materialization of principles, we propose a reflection on their use as a potential preamble to future AI legal frameworks, since an increasing demand in national and international regulations has emerged following the gains and drift of those ethical charters.

3 Building on Ethical Charters to Expand the AI Normative Landscape

The formalization of ethical charters should also be explored with respect to its potential to help develop AI regulations, either locally or internationally. Despite the valuable yet unequal contribution of ethical charters, the need to go beyond them – without discarding them – has been increasingly recognized; the next step to enriching the AI normative landscape is to propose additional binding norms. This is probably a natural development given that AI ethics grew from an underexplored area in

⁷ See example of the standards of practice proposed by IEEE and validated by the engineers concerned.

significant need for societal benchmarks into an increasingly studied and refined field. One of the reasons for this recognition relates to the previously underlined fact that ethical charters have shown their limits in governing some AI developments, including for powerful industries pushing their business interests in unwanted societal directions. This has triggered a demand for further regulations, eventually inspired by such charters (that is, they could serve as a compass for certain legal developments). Another reason is that these charters have made valuable, yet incomplete progress in defining the basis of an international common framework offering opportunities for normative scaling up in an area where boundaries are often erased. These two aspects can be considered as the next steps in AI governance, and are explored in the next sections.

3.1 The Deployment of AI Regulations

The recent shift from the demand for ethical to legal guidance is perhaps most strikingly embodied by Google CEO Sundar Pichai's January 2020 statement, in which he mentioned that AI needs to be regulated, especially in areas like self-driving cars and healthcare technology: "Regulation and self-regulation, via a code of ethics and an ethics board, might not be enough [...]" (BBC 2020)⁸ Some would say that, at this point, the demand for legislation has become so important that Google could not oppose it without significant public backlash; so it is better to be part of the regulatory process in order to shape its development.

The risk of not moving towards additional legal governance despite the current effervescence in ethical work is real and multifaceted. Stakeholders have started exposing the impact of underregulated AI, pointing to risks related to privacy protection, solidarity, responsibility, discrimination, health or inclusiveness issues. Well-known scandals, such as Cambridge Analytica and Buolamwini (2018),⁹ have helped bring much public exposure to these issues. During the public consultation process that led to the Montreal Declaration for Responsible AI Development, the demand for additional legal protection came up as the policy option most required by stakeholders.¹⁰ That being said, it would be inaccurate to state that there are currently no AI regulations; legal regimes around the world already have some tools to address responsibility issues, whether or not they specifically relate to AI. For example, AI generates a set of potential opportunities and problems that can be addressed, even if not always adequately, under usual legal regimes like tort law, contracts, criminal law,

⁸ <https://www.bbc.com/news/technology-51178198>; Other key technology players like Microsoft president have also called for further regulation Monica Nickelsburg, "Microsoft President Brad Smith Calls for AI Regulations at Davos" (21 January 2020), online: *GeekWire* <https://www.geekwire.com/2020/microsoft-president-brad-smith-calls-ai-regulation-davos/>.

⁹ The Guardian, "The Cambridge Analytica scandal changed the world – but it didn't change Facebook" <https://www.theguardian.com/technology/2019/mar/17/the-cambridge-analytica-scandal-changed-the-world-but-it-didnt-change-facebook>.

¹⁰ Report of the Montreal Declaration for the Responsible Development of Artificial Intelligence (2018). Part 6 - Priority areas and their recommendations for the responsible development of AI. https://5da05b0d-f158-4af28b9f892984c33739.filesusr.com/ugd/ebc3a3_d6a627b2f8644a30ae174762557da6fc.pdf.

administrative law, and so on.¹¹ However, the capacity of these regimes in the context of AI leaves some blind or weak spots for legal oversight or insufficiently tackles some of the novel issues that AI raises, such as consent in the Age of Big Data, lack of explicability and transparency in algorithms and unprecedented uses of AI technology like automated killer robots, a new form of political manipulation and facial recognition.¹²

If not addressed by suitable legal frameworks, such issues can erode trust in political institutions, researchers and innovators, to name a few. As a result, this situation can weaken the buy-in required from societal stakeholders for innovations to take hold. Such a lack of public acceptability will then create a significant chasm between development and implementation of an innovation (and, arguably, even more so for responsible innovation). Some authors (Panch et al. 2019) have called this gap between development and implementation - which is due to different factors - the most “inconvenient truth” about AI, as it is already noticed in certain sectors of AI activity. This chasm could result in significant lost opportunities for valuable AI developments likely to improve society, the environment and individual well-being, as well as financial and research effort loss due to unusable or underuse AI developments (Lovis 2019; Flood and Régis 2021). Furthermore, uncertainty regarding legal hazards, whether real or perceived, is in itself an obstacle to innovation. Relevant literature already mentions this impediment, and current research projects evaluating AI implementation in healthcare professionals’ practices highlights their resistance to innovation due to a lack of clarity with respect to their legal responsibility.¹³ There is thus an evident need to clarify legal frameworks and responsibilities in order to sustain innovation. In other words, legal predictability must be enhanced; this is a different goal than an ongoing process of ethical reflexivity. Sometimes, legal predictability will require the redesign of legal frameworks; other times, it will require increased knowledge transfer regarding their meaning in specific contexts. (Girard and Régis 2020).¹⁴

¹¹ See for example: Marion Oswald, Algorithm-assisted decision-making in the public sector: framing the issues using administrative law rules governing discretionary power, (2018), *Phil. Trans. R. Soc. A.* 376; Michael Froomkin, Ian Kerr and Joelle Pineau, «When AIs Outperform Doctors: Confronting the Challenges of a Tort-Induced Over-Reliance on Machine Learning, (2019) vol 61:33, *Arizona Law Review*, 33; Frank Pasquale, Data-Informed Duties in AI development, *Columbia Law Review* 119:1917 (2019).

¹² For examples see the different essays in Glenn Cohen and al. *Big Data, Health Law, and Bioethics*, 2018 Cambridge University Press; Ian Kerr & Katie Szilagyi, “Asleep at the switch? How killer robots become a force multiplier of military necessity” in Ryan M. Calo, Michael Froomkin, Ian Kerr, eds, *Robot Law* (Cheltenham: Edward Elgar Pub Ltd, 2016) 333; Harry Surden, 35 *Ga. St. U. L. Rev.* 1305 (2018–2019) *Artificial Intelligence and Law: An Overview*.

¹³ See OMS, https://apps.who.int/iris/bitstream/handle/10665/44497/9789241564144_eng.pdf;jsessionid=B511E3B52CFC230B4B4B4CB309F7B4F8?sequence=1, at p. 11.

¹⁴ Indeed, sometimes, the legal framework is clear, but its understanding by professionals is not; it then becomes a matter of better communicating the meaning of such frameworks. See for example Marie-Andrée Girard and Catherine Régis, « La collaboration interprofessionnelle: une pratique complexe dans un environnement juridique tout aussi complexe » 2020 *J.D.S.A.M.* numéro 25 at 153.

Yet, legal developments in innovation logic may be in the midst of a paradigm freeze, encapsulated in Munro's quote: "In the early days of emerging technology, we have power but insufficient clarity to act. In later days, we have more clarity, but declining power."¹⁵ Considering the procedural requirements of legislative processes and the crystallizing, constraining effect of law on conduct, proceeding with legal developments usually requires a certain maturity in understanding social phenomena. But when such an understanding is finally acquired, it can become more difficult to regulate due to the resources and interests invested in innovation.

AI offers a good illustration of this "paradigm freeze" situation (Kuhn 1983). Technology, especially machine learning, has evolved at an incredibly rapid pace. AI progress is now considered as following a rapid obsolescence cycle, in which the technology will significantly evolve every few years. In addition, machine learning is a complex mathematical and computer technique, one that requires a significant learning curve for many non-specialists, including regulators. During the early stages of development and implementation, it was therefore difficult (and perhaps still is) to identify precisely the aspects of this technique that need to be regulated, and why and how. Understanding the functioning of AI and its risks, benefits and challenges has significantly evolved thanks to the ethics community and ethical charters. When these concepts became more familiar, a massive amount of investment was injected to support this innovation.¹⁶ It then became more and more difficult for governments around the world to formally regulate AI, at least in a way that could limit its development, for instance to better balance innovation with stronger digital rights for consumers, patients and so on.¹⁷ As Boisson de Chazournes (2014) stresses, in the digital world, the repeated actions of individuals can eventually shape norms: behaviour influences normativity. This can create a path of dependency in regulation, where innovation becomes normative and law eventually enshrines, at least partially, what becomes "normal practice" through people's increasingly routine use of technology. Put differently, digital technology strongly and rapidly affects people's social expectations. This is another reason why the impact of ethical charters has been important to the extent that they influence technology design, as they could later contribute to shaping, at least to some degree, such expectations.

The ethical charters developed so far have offered some promise to overcome, at least partially, this potential paradigm freeze. When law could not move, ethics was active in targeting priority areas of normative action (sometimes through public consultation processes), democratizing access to AI information and bringing to the attention of policy-makers (in a way, making it almost unavoidable for them to stay idle) the importance of interventionist measures to balance innovation dynamics with

¹⁵ See: <https://www.danmunro.ca/blog/2019/1/16/risk-uncertainty-and-the-governance-dilemma-for-artificial-intelligence>.

¹⁶ See: AI Index Report de 2019. https://hai.stanford.edu/sites/default/files/ai_index_2019_report.pdf.

¹⁷ That being said, law is more than a tool to distribute the costs of incidents among actors developing and deploying innovation; it can contribute to share the benefits of innovation, for example by influencing private industries development and labor conditions.

developers' responsible actions. The law can now build upon these valuable ethical assets to better target regulatory options. (Petitgand and Régis 2019).¹⁸

The exercise of building legal developments like legislation on some ethical charters will necessarily require adjustments. The law requires a capacity to operationalize its normative postulates (to make them enforceable) which does not always concern ethics, as mentioned in the first part of this paper. Yet, the formalization of ethics will contribute to providing an “applicability test” that could be helpful at this point, especially if it includes a feedback process allowing stakeholders to further understand what works and what is not normatively wise. While certain ethical principles can be more easily translated into legal changes, others are simply impractical. For example, a principle of *justice* (at least without further precision) does not correlate with actionable legal duties that can be imposed on stakeholders, whereas principles like *respect for human autonomy* and *explicability* do so more easily, as they are already part of legal corpora in many legal systems. Besides, some ethical guidelines are more granular than others, offering richer normative content to explore legal reforms. The Montreal Declaration is a relevant example with ten main principles, many sub-principles and a side report detailing policy recommendations.¹⁹ In fact, some legal researchers have already started to build on ethical charters to propose legal reforms or legal interpretation. For example, some authors (Lutun 2019; Régis 2019) have recommended enshrining a patient's right to consent (and not just to be informed) to physicians' use of AI tools based on Principle 9 of the Montreal Declaration.²⁰ This work will arguably need to continue even if the law does not ultimately depend on ethical guidelines to evolve. Despite possible “lost in translation” problems between law and ethics in developing regulation, the intersection of the two fields provides normative guidance and opportunities that, if mobilized appropriately, can amplify each domain's contribution to AI. This interaction of law and ethics has already been mobilized before, including in the field of medicine. The same could be done in the field of AI, despite the undeniable challenges that such a widely encompassing, mostly private and fast-developing culture of new professionalism entails (Mittelstadt 2019b).

At the end of the day, while the roles of law and ethics should not be confused, it would be a mistake to position the two as necessarily complementing its responses to

¹⁸ For example of such work: a) High-Level panel on AI de la Commission européennes: <https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top> and <https://ec.europa.eu/digital-single-market/en/news/policy-and-investment-recommendations-trustworthy-artificial-intelligence>; b) Montreal Declaration for the Responsible Development of Artificial Intelligence. Journal de Droit de la Santé et de l'Assurance Maladie (Journal of Health and Health Insurance Law).

¹⁹ www.montrealdeclaration-responsibleai.com.

²⁰ Adele Lutun (2019), « L'article 11 du projet de loi bioéthique français prend-il en compte les principes de la Déclaration de Montréal pour un développement responsable de l'intelligence artificielle ? / Does Art. 11 of the French draft bioethics law take into account the principles of the Montreal Declaration on the Responsible Development of Artificial Intelligence? Available online: <https://www.chairesante.ca/articles/2019/larticle-11-du-projet-de-loi-bioethique-francais-prend-il-en-compte-les-principes-de-la-declaration-de-montreal-pour-un-developpement-responsable-de-lintelligence-artificielle/>»; Catherine Régis, « Perspectives internationales sur la régulation de l'IA dans le domaine de la santé », conference given for the Entretiens Droit & Santé, Université Paris Descartes, Paris, 6 décembre 2019.

the “flaws” of the other. These two normative approaches have different contributions and logic that are both valuable. Acknowledging these distinctions can allow the development of a coherent normative strategy for responsible AI innovation that embraces the complementarity and prospect of each; neither of these approaches is a standalone instrument to address entirely the complex field of AI. Even if the legal field deploys further regulations to address AI issues, ethics will still be required. Essentially based on general enforceable rules that apply to everyone, law will not have the means to address every context-specific variation and the rapid development that AI will trigger. Due to their less formal and fast-paced development process (at least compared with law), ethics and ethical charters need to stay responsive, relevant and purposeful in the AI normative landscape if they are to add agility and layers of reflexivity in AI normativity. The complementarity of the two types of norms forms a rich juncture in moving towards the next steps in AI normative framework development, including at the international level.

3.2 Seizing Opportunities for an International Scaling-Up of AI Normativity

Ethical charters could also be a valuable asset to expand the normative landscape at the international level. International norms often start with the identification of a global problem that no country alone can adequately address or that touches on issues with a reach beyond territorial borders, for instance, due to their impact on human dignity and life, global economy and the environment. The United Nations’ High-level Panel on Digital Cooperation highlights the need to develop collaborative international governance tools to manage Big Data and AI, considering the undeniable effect that such digital products will have on humanity.²¹ Moreover, the scope of many AI issues extends well beyond national borders, with digital data travelling from one country to another, algorithms having an impact on worldwide conducts, human life and the environment, and cyberattacks disturbing organizations, governments, universities, hospitals and other institutions around the world. Facial recognition data collected on travellers in airports, cookies gathering information on cyber-consumers, labour transformations due to AI progress, digital biosurveillance tools detecting infectious disease outbreaks and the 2017 WannaCry cyberattack with significant impact on organizations across continents are but a few examples of such impact (Martin 2019; Alford 2019).²² There is thus a strong argument to be made in favour of scaling-up normative work in AI, in order to develop international norms. This is also in line, at least partially, with UNESCO’s very recent work: an international group of experts has started drafting a global recommendation on the ethics of AI²³.

²¹ See: United-Nations, *The Age of Digital Interdependence*, Report of the UN Secretary-General’s High-Level on Digital Cooperation, 2019.

²² CNN.com (2019). When seeing is no longer believing - Inside the Pentagon’s race against deepfake videos. Online: <https://www.cnn.com/interactive/2019/01/business/pentagons-race-against-deepfakes/>.

²³ <https://en.unesco.org/news/unescos-expert-group-revises-draft-text-recommendation-ethics-artificial-intelligence> (date of access: 2 december 2020).

As a contribution to this process, all the ethical charters developed worldwide have started to identify points of convergence and divergence in AI governance. These charters have also demonstrated that many AI issues and their potential responses are global in nature. Building international consensus takes time and resources, particularly when developing treaties in public international law. As mentioned earlier, the work of Jobin et al. (2019), Zeng (2019) and Fjeld et al. (2020) is instrumental in this regard, as they mapped and analyzed the voluminous body of ethical AI charters and guidelines. Their contribution can serve as the starting point of a reflection process on future global digital rights, given the analysis and synthesis effort already exerted. For instance, Jobin et al. identified five ethical principles that demonstrate clearer international convergence: transparency, justice and fairness, non-maleficence, responsibility and privacy. However, the research team noted that some of these principles do not necessarily share the same definition. Nonetheless, these points of convergence provide a good start to identifying the shared principles that could eventually become a consensus for international norms, referred to as the “normative core” by Fjeld, J. *et al.* This is indeed a good start – and surely not a final endeavour – since, as mentioned earlier, some of the principles do not fully integrate ethical requirements, underrepresent some important voices and might be culturally specific (Jobin et al. 2019). The concept of justice underlying the charters is also indicative of how moral issues and moral reasoning are addressed. In fact, the concept of justice has been criticized in feminist studies of the 1980s in the field of moral development.²⁴ Hagendorff (2020) has highlighted the biases that continue to be perpetuated in his study of the 22 charters by pointing out the fact that the AI ethics discourse is primarily shaped by men and a way of approaching moral problems. (2020, p. 103). These limitations must be kept in mind, and compensated for, if they are to be exploited for future comprehensive and inclusive normative work.

While ethical AI guidelines could be helpful in accelerating international AI framework developments, many questions have yet to be answered regarding how this international work should be organized, including the choice of organization to lead the process and whether it should be ethically or legally driven (or both). The leadership role is key, as it will influence how the framework will be initiated and how it will later unfold. Identifying the best international organization to hold such a role, either exclusively or in partnership with other stakeholders, is open for debate. Interesting international AI partnerships are currently being established to accelerate global debates and actions likely to contribute to the process. Establishing international partnerships are often the outcome of initiatives that fostered discussions on ethical principles and the importance of maintaining and integrating ethical and legal benchmarks into the process to implement governance. The Global Partnership on Responsible Artificial Intelligence (GPAI) is a joint France-Canada initiative that has to

²⁴ See: Carol Gilligan, *In a Different Voice: Psychological Theory and Women's Development* (Cambridge, MA: Harvard University Press, 1982). Eva Feder Kittay and Diana T. Meyers, eds., *Women and Moral Theory* (Totowa, NJ: Rowman & Littlefield, 1987). Neil Noddings, *Caring: A Feminine Approach to Ethics and Moral Education* (Berkeley, CA: University of California Press, 1984). Blum, Lawrence A., 1988, “Gilligan and Kohlberg: Implications for Moral Theory,” *Ethics*, 98 (3): 472–491. Naussbaum 1999, *Sex and Social Justice*, Oxford: Oxford University Press.

date brought together various countries (such as Australia, Canada, France, Germany, Italy, Japan, Korea, Mexico, the United Kingdom, the United-States and New Zealand); this is a fine example at least in theory, of international governance. That being said, this initiative is too recent to fully appreciate its impact.

However, we would argue that there is need for a leadership role from an international intergovernmental organization like the United Nations (UN), for at least two reasons. First, considering the major economic interest in AI development and deployment, most of which private, an international dialogue needs to be organized through an institution that is relatively remote from private industries (even if they are included in the dialogue process), that can gather top-level officials from different countries (and thus top-level government commitment and eventual involvement of regulatory powers) and that represents the global public interest. Second, international intergovernmental organizations can often mobilize public international law, as is clearly the case for the UN,²⁵ allowing them to foster the development of normative actions of different natures (binding or not) that guide States around the world. Of course, States remain involved in proposing and adopting such norms, but the UN can ultimately achieve something that no State alone can do. How the UN should eventually coordinate its AI normative actions with other intergovernmental organizations, sometimes attached to the UN itself (such as the World Health Organization, UNESCO and the International Labour Organisation) needs to be further examined. That being said, the UN is capable of engaging conversations around multi-sector and multi-stakeholder issues. This conversation is a necessity in AI, since it generates challenges for almost every sector of societal activity. The UN's role is therefore important to addressing AI from a transversal perspective, which will be essential for normative coherence: similar issues in education, health, environment, and so on should trigger similar normative responses.

The kind of normative instruments that the UN or other international intergovernmental or transnational organizations should develop to support “good” AI is surely a complex and requisite debate that is beyond the scope of this paper. There is a valid argument to make in favour of establishing at least some form of binding international AI norms, arguably an international treaty, acknowledging the points raised previously regarding the need for additional forms of enforceable AI regulation.²⁶ Whether or not they are binding, international norms can have an impact on the AI world, even if they do not address all the potential opportunities and problems that this thriving innovation will raise – for instance, public international law faces challenges on its own, such as its enforcement capabilities (Kantorowicz-Reznichenko (2020); Raustiala 2000; Hongju Koh 1997). Legal scholars have demonstrated for some time now that binding (hard law) and non-binding (soft law) both produce some effects in domestic law, even though these might be different and follow a different path of implementation into

²⁵ See the Charter of the United Nations: <https://treaties.un.org/doc/publication/ctc/uncharter.pdf>.

²⁶ The binding and non-binding distinction is a classic division in public international law. A norm is considered to be binding when it creates mandatory obligations for member states (hard law) and non-binding when it is intended to assist action, consultation, negotiation and cooperation without explicit obligations or specific adoption procedures (soft law).

domestic law (Shaffer and Pollack 2009; Hillgenberg 1999; Finnemore and Sikkink 1998; Betts and Orchard 2014; Gostin 2014; Weil 1983; Hathaway and Shapiro 2011).

As such, in the process of devising and developing international AI frameworks, a normative strategy should not be designed *in abstracto* from the very nature of a particular norm, but instead should analyse and evaluate the advantages and drawbacks specific to each norm by considering legal, political, economic and ethical concerns relating to the AI issues at stake (Régis and Kastler 2018). For instance, the process of adopting a norm both at an international and national level could favour choosing a non-binding instrument. The process of creating binding instruments is long, complex, rigid and costly. As a result, they are less suitable for issues that require a rapid response and flexibility to adapt to scientific developments or that concern only a limited number of states (Gostin 2014; Chevalier 1998). The process can also be prolonged by national implementation procedures required for binding law. As another illustration, the existence of “competing” binding legal norms could require the adoption of similar norms to gain political and legal traction. For instance, norms may have to be binding in contexts where they will be in tension with other powerful binding international norms, such as trade and intellectual property laws. (Régis and Kastler 2018) This will clearly be the case for AI at some point. The creation of such frameworks will also require first exploring how existing international norms (in treaties, declarations, resolutions, and so on) might already cover or leave gaps for AI challenges – even despite directly referencing AI or Big Data – in order to build a coherent and relevant normative response²⁷.

At the end of the day, the best international AI framework will probably be a skilful composite of complementary binding and non-binding norms forming a global strategy for responsible AI. The five shared principles found in ethical guidelines could again serve as a starting point, once stakeholders agree on their meaning and translate them into useable material for developing international law. And if such principles can be validated as triggering social acceptability, or even social preferability (Floridi and Taddeo 2016), from a global perspective – a key ethical consideration – the norms developed based on these principles will more likely induce voluntary compliance. This is an ideal position in international law, where state sovereignty sometimes comes in the way of enforceability.

4 Conclusion

The formalization of ethical charters is both a requisite and essential step to providing a more appropriate normative framework for artificial intelligence ethics and digital rights. However, this process must be assessed to measure the actual scope of its relevance. In this regard, a wealth of organizational ethics research is available, particularly on ethical institutionalization. It provides interesting insights on how to gain a better understanding of the factors that could foster the integration of ethical principles

²⁷ United-Nations, *The Age of Digital Interdependence*, Report of the UN Secretary-General’s High-Level on Digital Cooperation, 2019.

in AI systems. The declaration *Toward Trustworthy AI Development: Mechanisms for Supporting Verifiable Claims*²⁸ is the outcome of the international AI research community's study of the issue and another significant milestone in ethical formalization and institutionalization. The analysis clearly emphasizes AI drawbacks identified in the declaration yet acknowledges the importance of ethics. Nevertheless, the authors stress that a significant amount of work has yet to be done in formulating ethical principles before they are actually applied.

Overall, ethical charters are instruments to be used to establish a reflective ethical approach while highlighting priority values. These values can then serve as a compass for the development of other normative terms, such as legal and global governance standards. The challenge lies in developing an AI normative strategy through an integrated process, by carefully drawing on the strengths of various types of norms, whether they are binding or not, to establish a trustworthy AI governance. This challenge includes maintaining an interrelation, and, more broadly, a dialogue between ethics and law, as addressed in this article. However, to date this correlation has been insufficiently mobilized across the world of AI governance, and it represents one of the greatest challenges to ensuring the future well-being of humankind as AI innovations are deployed, globally, and begin to affect a wide range of aspects (even all) of human life.

References

- Adelman, H.: Morality and ethics in organizational administration. *J. Bus. Ethics* **10**(9), 665–678 (1991)
- Alford, J.: NHS cyber-attacks could delay life-saving care and cost millions. Imperial College London (2019). <https://www.imperial.ac.uk/news/193151/nhs-cyber-attacks-could-delay-life-saving-care/>.
- Bairaktarova, D., Woodcock, A.: Engineering student's ethical awareness and behavior: a new motivational model. *Sci. Eng. Ethics* **23**, 1129–1157 (2017)
- Barabas, C., Dinakar, K., Ito, J., Virza, M., Zittrain, J.: Interventions over predictions: reframing the ethical debate for actuarial risk assessment. In: *Proceedings of FAT* Conference (FAT* 2018)*, vol. 9. ACM, New York (2018). ISBN: 978-92-3-100363-9
- Bégin, L.: *L'éthique au travail*, éditions Liber, Québec (2009)
- Betts, A., Orchard, P.: Introduction: the normative institutionalization-implementation gap. In: Betts, A., Orchard, P. (eds.) *Implementation and World Politics: How International Norms Change Practice*, pp. 1–26. Oxford University Press, Oxford (2014)
- Bietti, E.: From ethics washing to ethics bashing: a view on tech ethics from within moral philosophy. In: *DRAFT – Final Paper Published in the Proceedings to ACM FAT* Conference (2019)*
- Blum, L.A.: Gilligan and Kohlberg: implications for moral theory. *Ethics* **98**(3), 472–491 (1988)
- Boisson de Chazournes, L.: Standards, régulation internationale et organisations internationales. In: Frydman, B., Van Waeyenberge, A. (eds.) *Gouverner par les standards et les indicateurs. De Hume aux rankings*, pp. 69–92. Bruylant/LGDJ, Bruxelles/Paris (2014)

²⁸ See *Toward Trustworthy AI Development: Mechanisms for Supporting Verifiable Claims*, a major charter created by a large portion of the international AI research community. <https://arxiv.org/abs/2004.07213>.

- Boisvert, Y.: *L'institutionnalisation de l'éthique gouvernementale: Quelle place pour l'éthique ?* Presses de l'Université du Québec (2011)
- Brent Mittelstadt, B.: Principles alone cannot guarantee ethical AI. *Nat. Mach. Intell.* **1**, 501–550 (2019)
- Buolamwini, J., Gebru, T.: Gender shades: intersectional accuracy disparities in commercial gender classification. In: *Conference on Fairness, Accountability and Transparency*, pp. 77–91 (2018)
- Carroll, A., Bucholtz, A.: *Business Society: Ethics and Stakeholder Management*, 4th edn. South-Western College Publishing, Cincinnati (1999)
- Chevallier, J.: Vers un droit post-moderne? Les transformations de la régulation juridique. *Revue du droit public et de la science politique en France et à l'étranger* **3**, 659–690 (1998)
- Cohen, G., et al.: *Big Data, Health Law, and Bioethics*. Cambridge University Press, Cambridge (2018)
- Dratwa, J.: *Dans quel monde voulons-nous vivre?*, vol. 4. éditions ISTE, London (2019)
- Flood, C.M., Régis, C.: Artificial intelligence and health law. In: Scassa, T., Martin-Bariteau, F. (eds.) *Artificial Intelligence and the Law in Canada*, Lexis Nexis (2021). (accepted for publication, to be published in 2021)
- Finnemore, M., Sikkink, K.: International norm dynamics and political change. *Int. Organ.* **52**(4), 887–917 (1998)
- Fjeld, J., Achten, N., Hilligoss, H., Nagy, H., Srikumar, M.: *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI*. Berkman Klein Center for Internet & Society (2020). <https://dash.harvard.edu/handle/1/42160420>
- Floridi, L.: Translating principles into practices of digital ethics: five risks of being unethical. *Philos. Technol.* **32**(2), 185–193 (2019)
- Floridi, L., Taddeo, M.: What is data ethics? *Philos. Trans. R. Soc. A* **374**(2083) (2016). Available at SSRN: <https://ssrn.com/abstract=2907744>
- Fontanel, J., Sushcheva, N.: *La puissance des GAFAM: Réalités, apports et dangers*. *Annuaire français de relations internationales*, La Documentation française, Paris (2019). XX. fhal02196915
- Froomkin, M., Kerr, I., Pineau, J.: When AIs outperform doctors: confronting the challenges of a tort-induced over-reliance on machine learning. *Univ. Miami Legal Stud. Res. Paper* **61**, 33 (2019)
- Gates, M., Ma, J. (eds.): *The age of digital interdependence*. Report of the UN Secretary-General's High-level Panel on Digital Cooperation (2019)
- Gautrais, V.: *Neutralité technologique: rédaction et interprétation des lois face aux technologies*. Thémis, Montréal (2012)
- Ghafur, S., Kristensen, S., Honeyford, K., et al.: A retrospective impact analysis of the WannaCry cyberattack on the NHS. *NPJ Digit. Med.* **2**, 98–105 (2019)
- Gilligan, C.: *In a Different Voice: Psychological Theory and Women's Development*. Harvard University Press, Cambridge (1982)
- Girard, M.-A., Régis, C.: La collaboration interprofessionnelle: une pratique complexe dans un environnement juridique tout aussi complexe. *J.D.S.A.M.* **25**, 153 (2020)
- Google boss Sundar Pichai calls for AI regulation (2020, 20 janvier). *BBC News* (2020). <https://www.bbc.com/news/technology-51178198>
- Gostin, L.: *Global health law*, pp. 64–67. Harvard University Press, Cambridge (2014)
- Green, B.: "Good" isn't good enough", Neurips conference paper (2019). https://aiforsocialgood.github.io/neurips2019/accepted/track3/pdfs/67_aaisg_neurips2019.pdf
- Hagendorff, T.: The ethics of AI ethics: an evaluation of guidelines. *Mind. Mach.* **30**, 99–120 (2020). <https://doi.org/10.1007/s11023-020-09517-8>

- Hathaway, O., Shapiro Scott, K.: Outcasting: enforcement in domestic and international law. *Yale Law J.* **121**, 252–349 (2011)
- Hillgenberg, H.: A fresh look at soft law. *Eur. J. Int. Law* **10**, 499–515 (1999)
- Finnemore, M., Sikkink, K.: International norm dynamics and political change. *Int. Organ.* **52**, 887–917 (1999)
- Hongju Koh, H.: Bringing international law home. *Houston Law Rev.* **35**, 623–681 (1997)
- <https://www.forbes.com/sites/nicolemartin1/2019/09/25/the-major-concerns-around-facial-recognition-technology/#558f9ba94fe3>, Justine Alford (2019) <https://www.imperial.ac.uk/news/193151/nhs-cyber-attacks-could-delay-life-saving-care/>, <https://www.nature.com/articles/s41746-019-0161-6>
- Jobin, A., Ienca, M., Vayena, E.: The global landscape of AI ethics guidelines. *Nat. Mach. Intell.* **1**, 389–399 (2019)
- Kantorowicz-Reznichenko, E.: The enforcement of international law. In: Hosli, M., Selleslaghs, J. (eds.) *The Changing Global Order. United Nations University Series on Regionalism*, vol. 17. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-21603-0_18
- Kerr, I., Szilagy, K.: Asleep at the switch? How killer robots become a force multiplier of military necessity. In: Calo, R., Froomkin, M., Kerr, I. (eds.) *Robot Law*, pp. 333–366. Edward Elgar Pub Ltd., Cheltenham (2016)
- Kittay, E.F., Diana, T.M. (eds.): *Women and Moral Theory*. Rowman & Littlefield, Totowa (1987)
- Koh, H.: Why do nations obey international law? *Yale Law J.* **106**, 2599–2659 (1997)
- Kuhn, T.: *La Structure des révolutions scientifiques*, Flammarion, coll. «Champs», Montréal (1983)
- Langlois, L.: Le leadership éthique: un mode de gouvernance responsabilisant, dans *Cinq questions d'éthique organisationnelle*, sous la direction de Luc Bégin, éditions Nota Bene, pp. 139–172 (2014)
- Lovis, C.: Unlocking the power of artificial intelligence and big data in medicine. *J. Med. Internet Res.* **21**, 16607 (2019)
- Lutun, A.: L'article 11 du projet de loi bioéthique français prend-il en compte les principes de la Déclaration de Montréal pour un développement responsable de l'intelligence artificielle? (2019). <https://www.chairesante.ca/articles/2019/>
- Malherbe, J-F.: *Le nomade polyglotte : l'excellence éthique en postmodernité* (2000). éditions FIDES
- Martin, N.: *The Major Concerns Around Facial Recognition Technologies*. Forbes.com, 25 September 2019
- Mercier, S.: L'éthique dans les entreprises, Repères, no. 263, La Découverte & Syros (1999)
- Mercier, S.: L'éthique dans les entreprises, collection Repères, La Découverte (2004)
- Metzinger, T.: EU Guidelines: Ethics washing made in Europe. *Der Tagesspiegel* (2019). <https://www.tagesspiegel.de/politik/eu-guidelines-ethics-washing-made-in-europe/24195496.html>
- Miller, C., Coldicott, R.: People, power and technology: the tech workers' view (2019). <https://doteveryone.org.uk/report/workersview/>
- Mittelstadt, B.: Principles alone cannot guarantee ethical AI. *Nat. Mach. Intell.* **1**, 501–507 (2019)
- Montreal Declaration for Responsible AI. <https://www.montrealdeclaration-responsibleai.com>. https://5da05b0d-f158-4af2-8b9f892984c33739.filesusr.com/ugd/ebc3a3_d6a627b2f8644a30ae174762557da6fc.pdf
- Munro, D.: Risk, Uncertainty and the Governance Dilemma for Artificial Intelligence. Dan's blog (2019). <https://www.danmunro.ca/blog/2019/1/16/risk-uncertainty-and-the-governance-dilemma-for-artificial-intelligence>
- Murphy, P.: Creating Ethical Corporate Structures. *Sloan Manage. Rev.* **30**(2), 81–87 (1989)

- Naussbaum, M.: *Sex and Social Justice*. Oxford University Press, Oxford (1999)
- Noddings, N.: *Caring: A Feminine Approach to Ethics and Moral Education*. University of California Press, Berkeley (1984)
- Nye, D.: *Technology Matters: Questions to Live With*, vol. 47. MIT Press, Cambridge (2007)
- O’Neil, C.: *Weapons of math destruction: how big data increases inequality and threatens democracy*, 1st edn., p. 272. Crown, New York (2016)
- OMS Telemedicine Opportunities and developments in Member States, p. 11 (2010). https://apps.who.int/iris/bitstream/handle/10665/44497/9789241564144_eng.pdf;jsessionid=B511E3B52CFC230B4B4B4CB309F7B4F8?sequence=1
- Oswald, M.: Algorithm-assisted decision-making in the public sector: framing the issues using administrative law rules governing discretionary power. *Philos. Trans. A: Math. Phys. Eng. Sci.* **376**(2128), 20170359 (2018)
- Panch, T., Mattie, H., Celim, L.: The “inconvenient truth” about AI in healthcare. *Npj Digit. Med.* **2**, 1–3 (2019)
- Pasquale, F.: Data-Informed Duties in AI development. *Columbia Law Rev.* **14**, 1917–1940 (2019)
- Perrault, R., Shoham, Y., Brynjolfsson, E., et al.: Artificial intelligence index. 2019 annual report. Human-Centered Artificial Intelligence, Stanford (2019). https://hai.stanford.edu/sites/default/files/ai_index_2019_report.pdf
- Petitgand, C., Régis, C.: Principes éthiques et encadrement juridique de l’intelligence artificielle en santé : Exemple de la Déclaration de Montréal pour un développement responsable de l’intelligence artificielle. *Journal de Droit de la Santé et de l’Assurance Maladie* **22**, 102–106 (2019)
- Rapport de la Déclaration de Montréal pour un développement responsable de l’intelligence artificielle (2018) Partie 6 - Les chantiers prioritaires et leurs recommandations pour le développement responsable de l’IA. https://5da05b0d-f158-4af2-8b9f-892984c33739.filesusr.com/ugd/ebc3a3_d6a627b2f8644a30ae174762557da6fc.pdf
- Raustiala, K.: Compliance and effectiveness in international regulatory cooperation. *Case W. Reserve J. Int. Law* **32**, 387–440 (2000)
- Régis, C., Kastler, F.: Improving the world health organization’s normative strategy with respect to global health goals: what should we aim for? *Revue Belge de droit international*, p. 138 (2018)
- Reynaud, J.-D.: La régulation sociale, Dans *Le travail: autres réalités, autres regards*, *Revue internationale d’action communautaire* Numéro 25, Printemps, pp. 121–126 (1991)
- Salomon, R.C.: *The Little Philosophy Book*. Oxford Press, October 2007. ISBN: 9780195311143
- Shaffer, G., Pollack, M.: Hard vs soft law: alternatives, complements, and antagonists in international governance. *Minnesota Law Rev.* **94**, 706–799 (2009)
- Surden, H.: Artificial intelligence and law: an overview. *Georgia State Univ. Law Rev.* **35**, 1305–1337 (2019)
- The Charter of the United Nations, San Francisco (1945). <https://treaties.un.org/doc/publication/ctc/uncharter.pdf>
- The European Commission: Independent high-level expert group on artificial intelligence SET UP. European Commission (2019). Brussels. <https://ec.europa.eu/digital-single-market/en/news/policy-and-investment-recommendations-trustworthy-artificial-intelligence>
- United-Nations: The Age of Digital Interdependence, Report of the UN Secretary-General’s High-Level on Digital Cooperation (2019)
- Vold, K.V., Peters, D., Robinson, D., Calvo, R.: Responsible AI – two frameworks for ethical design practice. *IEEE Trans. Technol. Soc.* (2020). <https://doi.org/10.1109/TTS.2020.2974991>

- Villani, C.: Intelligence artificielle et travail, Rapport à la ministre du Travail et au secrétaire d'État auprès du premier ministre, chargé du numérique, France stratégie, mars 2018
- Weil, P.: Towards relative normativity in international law? *Am. J. Int. Law* **77**, 413–442 (1983)
- Xianhong, H., Neupane, B., Echaiz, L.F., Sibal, P., Rivera Lam, M.: Steering AI and advanced ICTs for knowledge societies: a Rights, Access, and Multi-stakeholder Perspective. UNESCO Series on Internet Freedom, Openness (2019)
- Zeng, Y., Lu, E., Huangfu, C.: Linking artificial intelligence principles, p. 14. arXiv:1 (2019)