# What's Next for AI Ethics, Policy, and Governance? A Global Overview

Daniel Schiff
Georgia Institute of Technology
Atlanta, Georgia

Justin Biddle
Georgia Institute of Technology
Atlanta, Georgia

Jason Borenstein
Georgia Institute of Technology
Atlanta, Georgia

Kelly Laas
Illinois Institute of Technology
Chicago, Illinois

## ABSTRACT

Since 2016, more than 80 AI ethics documents – including codes, principles, frameworks, and policy strategies – have been produced by corporations, governments, and NGOs. In this paper, we examine three topics of importance related to our ongoing empirical study of ethics and policy issues in these emerging documents. First, we review possible challenges associated with the relative homogeneity of the documents' creators. Second, we provide a novel typology of motivations to characterize both obvious and less obvious goals of the documents. Third, we discuss the varied impacts these documents may have on the AI governance landscape, including what factors are relevant to assessing whether a given document is likely to be successful in achieving its goals.

## CCS CONCEPTS

• **Computing methodologies → Artificial intelligence**; • **Social and professional topics → Codes of ethics**; **Government technology policy**.

## KEYWORDS

AI ethics, AI policy, corporate social responsibility

Major advances in artificial intelligence (AI) this decade have spurred intense interest by governments, companies, non-governmental organizations (NGOs), and the public. Advocates note the potential for vast economic growth and social benefit, while critics raise concerns about a collection of ethical, legal, and social risks. Such risks include algorithmic bias, disproportionate harms to vulnerable populations, failures of accountability and transparency, technological

unemployment—and even the possibility of existential threats from "superintelligence."

In response to this wave of innovation and attention, since 2016, governments, private sector organizations, and NGOs have rapidly begun to produce normative documents addressing AI [4]. These documents include ethics codes, principles, guidelines, frameworks, and policy strategies. To date, more than 80 English language documents exist [7]. At this early stage in the discourse of AI ethics and policy, these documents provide a rich opportunity to explore substantive concerns and strategies, emerging consensus and differences, motivations, and prospects for the future of AI governance.

While the explosion of interest and information may offer an important opportunity for AI stakeholders, it is also essential to approach the topic of global AI governance thoughtfully and systematically. AI includes many technologies that span a number of sectors, and different stakeholders disagree on the types and extent of governance that are appropriate. Some favor corporate self-regulation or collective industry regulation. They usually argue that governments lack the flexibility or understanding to regulate effectively, or that premature regulation would stifle innovation and competitiveness. Others believe that sector-specific laws or general AI regulation should be pursued.

Further, the AI-related documents themselves exhibit significant diversity in terms of their topical scope and length. These documents range from a single page code of ethics to policy strategies that run over a hundred pages. Some articulate research, innovation, or industry roadmaps, while others emphasize policy or discuss critical social and ethical issues.

Our aim, as an interdisciplinary team including a policy scholar, two philosophers, and an ethics code librarian, is to identify trends in the global AI ethics and governance conversation and examine important questions raised by these ethics documents. Within this document set, our team is seeking to identify the main ethical themes discussed by each document's authors as well as the policy sectors that are identified as most relevant to the AI domain going forward. Our research is substantively – though not exclusively – focused on ethics-related issues embedded within the document collection (henceforth "ethics documents").

We examined 88 ethics documents published between 2016 (the beginning point for most documents in this realm) to July 2019, including documents published by actors from the public, private, and NGO sectors. Two coders (of a total of four) are in the process of examining each document to assess the presence or absence of approximately 25 ethics topics and 15 policy sectors. Additionally,

we are evaluating further topics of interest, such as the degree of participatory engagement in the document's creation and the level of discussion of laws and regulations. At the time of this writing, our analysis is still ongoing.

This paper draws on our study to-date by exploring three issues and associated implications relevant to AI practitioners, scholars, policymakers, and other stakeholders: 1) Who is creating these ethics documents? 2) Why are they being produced? and 3) What impacts might these documents have on global AI governance? Understanding these issues is a step towards answering key questions about the future of formal and informal AI governance.

## 1 WHO IS LEADING THE CHARGE?

The discussion and debate surrounding AI policy and ethics is lively, taking place in board rooms and government offices, at international conferences and workshops, and in physical and online classrooms. However, a few entities have taken the lead in developing and publicizing AI ethics codes, principles, frameworks, and policy strategies.

### 1.1 Public Sector

Since 2016, public sector actors such as government agencies have released the largest number of AI-related ethics documents (n=50), more than half in our sample. The creators include European Union nations (especially France, Germany, and Italy), the United States, the United Kingdom, Canada, countries in the Nordic-Baltic region (such as Finland, Sweden, and Denmark), Japan, China, India, Mexico, Australia, and New Zealand. At the intergovernmental level, multiple bodies under the European Union have been active along with the OECD, the G7, and the G20. Many of documents produced by public sector actors are not solely ethics documents, but broader policy strategies that include significant ethics discussions.

While there is some regional diversity in this set, the documents' authors are primarily entities in the Global North along with Western-dominated intergovernmental institutions and a few major emerging economies (i.e., China, India, and Mexico). Other leading AI countries include Singapore, the Republic of Korea, Poland, Malta, Qatar, and Tunisia, though some have only begun engaging in conversation and not all have formalized their strategic thinking into public documents or policy action.

This apparent dominance by wealthy countries raises several questions and concerns. First, will low- and middle-income countries be represented adequately in global AI ethics and policy discourse either in terms of being invited to the table or having the resources to coordinate their own initiatives? Many wealthy countries — as service economies — are focused on local policy issues relevant to their contexts, such as healthcare, education, entertainment, and finance. Lack of attention to policy issues important in the Global South such as agriculture, tourism, and basic infrastructure could result in a lack of research and investment in AI solutions in these realms.

The countries who have a voice thus far are also angling for global economic competitiveness in their AI strategies. In the case of the United States and China, this means continued AI research and investment dominance, while countries such as the UK, Canada, and Denmark have identified special niches, such as ethical AI, to help with national branding and product differentiation. However, there is a risk that AI-driven growth defined and dominated by wealthy countries could detrimentally impact poorer ones.

The continued competition to recruit high-skill AI researchers suggests that brain drain will remain a problem for developing countries. Similarly, competitive efforts by wealthy countries to implement either labor-saving or capital-saving automation technologies in developing countries may not be in the latter's best interest. In its national AI strategy, *#AIFORALL*, India notes that its customer service and technical support workers may be displaced as their work is automated and brought back to wealthier countries, while Mexico's AI strategy notes the same for its manufacturing workers [9, 16]. Even more concerning, automating the work of teachers or healthcare workers via telehealth or intelligent tutoring systems could reduce investment in those countries' educational and health infrastructures, foreclosing on more appropriate development pathways in the interest of short-term improvements or profit.

To the extent that global AI and policy discourse is formulated based on a limited set of national or even supra-regional (i.e., the European Union) perspectives, it risks placing the benefit of those parties above that of the regions and communities most vulnerable around the world. Under this possible future, global inequalities and power asymmetries could be magnified, despite the attention to ethical and social issues emphasized by the leading AI players.

### 1.2 Private Sector

The private sector (n=20) produces about a quarter of the ethics documents we found. In the private sector, the pattern of AI leadership appears similar to that of the public sector. The private organizations producing AI ethics documents or statements are typically large and influential multinational corporations in the technology sector, including Microsoft, Google, IBM, Baidu, Tencent, Intel, Sony, Workday, SAP, and Sage. These corporations, along with a few others, are responsible for many of the world's major online, software, and hardware platforms, manage massive amounts of data, and lead AI research and development in several areas.

These corporations discuss standards for professional conduct (such as ethical principles for developers), internal governance strategies (such as advisory boards), and the role of external governance, sometimes promoting self-regulation or collective industry regulation while cautioning against government regulation they view as premature [11]. Some of these recommendations could shape and even narrow the range of normative policy responses, such as expanding anti-trust efforts or changing the privacy and data governance regime.

As in the public sector, there is a risk that small and medium businesses, consumers, and even governments face a power imbalance against these multinationals. Compounding network and platform effects and the sheer quantity of data favor an increased concentration of power for these few organizations. As noted in the UK's AI strategy, smaller-to-medium enterprises (SMEs), including startups, and even large corporations outside of these privileged few may have less access and therefore less opportunity to shape AI discourse and policy to their advantage [5].

## 1.3 NGO Sector

The NGO sector (n=18) exhibits the most diversity in terms of the types of entities that produced a document. This category includes both established organizations and organizations newly formed to focus explicitly on AI. There are professional associations like IEEE, The Japanese Society for Artificial Intelligence, and The Royal Society, think tanks and advocacy groups like the Future of Life Institute and the Partnership on AI, as well as documents formed through international workshops and collaborations like the *Asilomar AI Principles* and *The Montréal Declaration for Responsible AI.*

These documents constitute some of the most participatory in terms of who was involved in a document's development, sometimes with hundreds of participants and dozens of individual events featuring individuals from around the world and in different sectors. While some of the documents express deep ethical insight and breadth, they often stand in an outsider position, appealing to the private or public sectors to make change. In some cases these documents are written at an abstract level, without specific plans for operationalization or even a clear target audience.

This raises the question of to what extent these third-party organizations will be able to influence the private and public sector entities arguably responsible for most AI development, implementation, and governance decisions [1]. They may seek to influence public discourse and accrue signatories, but it remains to be seen whether these organizations have the power to shape global governance. It is notable, however, that many national AI strategies cite documents and experts or even solicit live testimony from this third sector.

## 2 SHINING LIGHT ON MOTIVATIONS

Organizations might develop an AI ethics document for a variety of reasons. For any given document, it can be difficult to determine which motivations are operative; stated motivations might not align with actual ones, and a given entity might be motivated by a mixture of factors simultaneously. Through our document analysis and discussions, we are developing a typology of six different motivations we believe characterize the entities who produce ethics documents. The six motivations are clustered into three pairs, each pair independent from the others. The first two motivations address end goals; the middle indicate target audiences, and the last focus on signaling. Despite the complexities inherent in unpacking underlying motives, a typology can serve as an analytic tool that may help examine the reasoning of entities behind AI ethics documents.

The first type of motivation is the motive of social responsibility — that is, the motive to promote social benefits and reduce risks of harm. We expect that many stakeholders, including longstanding advocacy and human rights groups, are genuinely motivated by a concern for the general well-being of society. One likely example is the IEEE's *Ethically Aligned Design*, a 294 page, multi-stakeholder document that includes extensive discussions of topics such as human rights, well-being, data agency, transparency, accountability, and prevention of misuse [20]. Similarly, the *OECD Principles on Artifical Intelligence* highlight the need for AI stakeholders to work proactively to promote "inclusive growth, sustainable development, and well-being" [17].

*2.0.1 Motivation of Competitive Advantage.* A second type of motivation is competitive advantage, including economic and political advantage. China's *New Generation of Artificial Intelligence Development Plan* describes AI as "the new focus of international competition," while Microsoft agrees that "companies and countries" that embrace AI "will fare best in the AI era" [11, 19]. Malta notes that its AI strategy maps its pursuit of "strategic competitive advantage in the global economy" while Mexico anticipates "an early mover advantage if it is among the first countries to announce an AI strategy" [9, 18].

However, a motive of competitive advantage does not imply the authors are not genuinely concerned about social responsibility. Competitive advantage can be part of a broader pro-social strategy, and securing economic growth and prosperity can itself be a good. These first two types of motivations are motivations to achieve particular ends, e.g., to promote social well-being or economic gain. In contrast, the next two motivations consider target audiences – internal or external.

*2.0.2 Motivation of Strategic Planning.* A third type of motivation is (internal) strategic planning or the production of an ethics document to support change within the organization itself. Such strategic planning includes how to incorporate ethical considerations into the organization's R&D, how to rethink public engagement, and how to market one's organization better. Other examples of this motive are general frameworks for pursuing AI, such as corporate ethics principles, and outlines of preliminary governance ideas. For example, the *U.S. National AI R&D Plan* recommends the development of "an AI R&D implementation framework... consistent with" their plan, while documents from France, Qatar, Mexico explicitly note their work is a blueprint or starting point for further strategic planning [15].

*2.0.3 Motivation of Strategic Intervention.* A fourth type of motivation is (external) strategic intervention—that is, the attempt to intervene in one's environment, including legal, political, economic, and social environments. For example, corporations and industry organizations may develop and adopt voluntary standards in the form of a code of ethics or a statement of principles to preempt regulations, thereby avoiding more restrictive laws being passed. Microsoft's [11] short book on AI suggests patience, arguing that "the most effective regulation can be achieved by providing all stakeholders with sufficient time to identify and articulate key principles." Additionally, ethics documents can impact external environments through the dissemination of narrative framings of AI ethics, promoting particular ethical concepts as being especially salient. Some documents focus on AI's potential for economic growth, while minimizing the possibility that this growth might exacerbate existing inequalities or harm vulnerable populations. How one frames ethics discussions around AI may therefore reflect decisions about strategic intervention.

While the first two types of motivations are, again, motivations to achieve particular ends, the second two can be seen as orthogonal strategies to affect changes that will help to achieve those ends. One could engage in strategic planning or strategic intervention for the purposes of promoting well-being or competitiveness (or both). The final pair of motivations concern signaling and perception of those signals.

*2.0.4 Motivation of Signaling Social Responsibility.* The fifth motivation is to signal social responsibility—that is, to appear to be socially responsible, whether or not one is actually motivated by genuine social responsibility. Again, this is orthogonal to the motive of genuine social responsibility. One could signal pro-social beliefs in order to counter misconceptions or be fairly recognized for their costly activities, such as subjecting themselves to increased scrutiny or lower economic gains. Success in fulfilling this goal can protect a company's or government's reputation, minimize risks of protest or boycott, or differentiate one's products or brand as ethical, responsible, or trustworthy. However, one could also signal responsibility to cover up for a lack of genuine action or even harms caused, a version of "ethics washing" [21]. We do not attribute this motivation to any particular stakeholder, but we are not the first to suggest that it could be operative.

*2.0.5 Motivation of Signaling Leadership.* The sixth is to signal leadership, whether or not one actually has a leadership role. Signaling leadership can help a country or organization be invited to the table, achieve access to expanded audiences or markets, or improve one's brand or international reputation. The recent rush to develop documents itself could be taken as a mark or signal of being a "player" in the realm. Notably, the strategic plans for Australia, India, Mexico, Tunisia, and New Zealand all cite the strategic plans of other countries and describe their own document as an entry point into this elite group. For example, India's [16] document observes that, "Governments in USA, UK, France, Japan and China have released their policy and strategy papers relating to AI. In order to establish a leadership role, it is important for India to take the plunge." Similarly, Finland states its goal to "be a global leader in applying artificial intelligence" [12]

Signaling leadership can be especially important for entities that are not already perceived as leaders. As Malta's [18] strategy argues, "the nature of AI development makes it possible for a single company or academic lab to put a small country on the map, even to lead and drive the agenda in one branch of AI." Mexico's [9] document emphasizes that it was "one of the first ten countries in the world to deliver a National Strategy for AI," and expresses pride to be "first nation in Latin America to join this elite club." Similarly, Malta's strategy discusses "positioning Malta… amongst the top 10 nations with the highest impact national AI programme," describing itself as "the Ultimate AI launchpad."

Again, this last pair of motivations is conceptually orthogonal to the ones above. One can signal leadership and responsibility while in fact leading responsibly. One can also signal leadership or responsibility in the interest of advancing ethical goals or competitive advantage (or both).

This typology is a first effort at characterizing the multilayered and varied motivations behind ethics documents and the organizations and authors producing them. While in our own study we lack a systematic method for inferring motivations, it may be possible to do so. This would involve disentangling stated from actual motivations as the documents vary in how explicitly they discuss motivations. The analysis would also involve deciphering tension between competing motivations for organizations, for documents produced by multiple authors, and even for multiple documents produced by the same organization. In short, though all of these

documents focus on ethics at least to some degree, the motivational aspects are far from simple.

## 3 HOW DO ETHICS DOCUMENTS CHANGE THE GLOBAL AI LANDSCAPE?

A final topic worthy of exploration is what impacts these documents might have on the development of global ethics and policy governance. The role of a given document may vary depending on the sector, region, state of policy-making in the external environment, and goals of the authors.

In some cases, the document is itself a concrete impact. A well-developed internal corporate strategy or government regulation may have practical utility or legal authority, and may spell out extant operational changes or new funding priorities. Alternatively, a document may exist largely to influence other organizations through persuasion, more typically found with NGO documents. As such, while many documents discuss similar topics, they may have more or less authority and capacity to produce change in a given sector or by a particular actor.

Also, because the role of a document could depend on motivations, it is not always easy to attribute success or failure to a document without understanding the authors' underlying motivations. A document may successfully improve one's brand without producing internal change or may influence public discourse without clear instantiation in regulation and law. Nevertheless, it is useful to review the possible impacts of ethics documents.

*3.0.1 Impacts on Other Ethics Documents.* It is clear that documents are already influencing one another across both sectors and regions. Numerous documents cite one other. Some provide extensive reviews of the issues and priorities identified in other countries, acquire testimony from experts, or even engage in site visits. To better understand these impacts, our ongoing research is attempting to identify the degree of consensus (or difference) across different organizations, regions, sectors, and over time. Thus far, dialogue between documents clearly serves as a way to both build on the ideas of others (learning or emulation) as well as to differentiate oneself (competition). In the former case, governments cite ethical themes developed from NGO-led participatory-style international AI conferences as inspiration. For example, documents developed by the European Commission, Australia, UK, and numerous others recognize documents like IEEE's *Ethically Aligned Design*, the *Asilomar AI Principles*, *The Montréal Declaration*, and the *Toronto Declaration*. In other words, the NGO sector is tangibly influencing ethics documents within the public sector.

Governments are also engaging with documents by learning how to best differentiate themselves, notably from both political opponents and allies. Malta's [18] strategy identifies that despite having a population of only half a million, it could "identify niche areas that put Malta at the forefront," in particular by serving as a regulatory sandbox. Denmark, the U.K, Canada, France, and others have identified ethical AI as a gap and opportunity for differentiation. Demark's *National Strategy for Artificial Intelligence* states that "Europe and Denmark should not copy the US or China. Both countries are investing heavily in artificial intelligence, but with little regard for responsibility, ethical principles and privacy" [14].

Therefore ethics documents can serve as examples of what to do, or indeed, what not to do. However, while there is already robust evidence of learning across documents, this is merely a preliminary form of impact. It remains to be seen if this cross-pollination will result in consensus or in more significant impacts.

*3.0.2 Internal Impacts.* A clear form of impact under an organization's control is internal organizational change, as might be expected for documents motivated by strategic planning. A government bureaucracy or private corporation could alter its practices by introducing new ethical review processes, tools for addressing algorithmic bias, or reporting practices to minimize excessive energy use. An organization could alter its governance through an internal advisory board or documentation of AI-related decision-making. Finally, an organization could alter its workforce hiring practices, composition, and skill set, such as by hiring a more diverse workforce to increase representativeness of marginalized groups or interdisciplinarity. In the private sector, Microsoft, Google, and SAP are all corporations that have articulated changes in business practices.

*3.0.3 External Impacts.* There are also numerous ways an ethics document can influence external environments. A document could influence corporate research and development practices in another entity (as opposed to one-self). It could impact the perception of customers or citizens by safeguarding or improving one's reputation. There is evidence that the documents are well aware of this possibility. For example, Denmark's national strategy proposes an ethics label for responsible businesses to help "consumers… choose the… responsible alternative." And as Finland's report on AI and work notes, "the values guiding companies in the development and use of artificial intelligence may in the future become highly important factors affecting companies' brand value. The leading companies in the sector have already understood this" [13]

*3.0.4 Public Discourse.* An ethics document could also influence public discourse more broadly by how it frames ethical issues, shaping which issues are seen as more (or less) important. Changes in discourse might minimize the visibility of certain problems associated with AI, like loss of privacy and labor displacement, or expand awareness of possible benefits, such as in the health sector. One example of framing is that some organizations argue that technical explainability of algorithmic outputs is not feasible or prudent, and instead emphasize process transparency and communication of less technical elements that stakeholders might understand better.

Another example of framing concerns inequality and social welfare. The majority of documents highlight risks of inequality and other harms caused by using AI to promote innovation and economic growth. However, very few talk about social welfare policies to address these problems. Even Finland [13], who notes in its report on AI and work that "responsibility for protecting those in the most vulnerable position must be shifted towards social policy," says that "making proposals related to social policy was not part of our working group's mandate." A frame which emphasizes growth at the expense of serious consideration of social welfare reform risks excluding these issues from policy discussions and the set of solutions that are ultimately considered.

*3.0.5 Policy.* Ethics documents can also inform (or constitute) policies, laws, and governance strategies. They might shape strategic funding priorities, national industrial policies, and research agendas. Both domestic and international governance are widely discussed. While this is arguably the most sought after change articulated in the documents, policy agenda setting, adoption, and diffusion are the result of a complex confluence of factors, of which ethical discussion in documents is only a part.

As noted previously, there is evidence of influence at the policy level, for example, illustrated by the U.S. National AI R&D Strategic Plan when it quotes the Future of Life as a source of inspiration. Similarly, Malta [18] states that the [3] Asilomar AI Principles "which promote ethics, values, privacy and the common good as core attributes will guide the Strategy that is being developed." Documents can therefore influence governance and policy through influencing political discourse, framing social problems, and identifying policy solutions. For this reason, our research seeks to assess which ethical issues are thought most prominent across documents, and what policy sectors are argued to be most impacted by AI.

*3.0.6 Stakeholder Ecosystem.* Ethics documents can also shape the broader stakeholder ecosystem. Numerous documents suggest strategies for reforming educational priorities and systems, such as by focusing on STEM skills, or alternatively social-emotional skills that are harder to computerize. This includes not just K-12 and higher education but also continuing education and life-long training. The funding and incentives behind academic research can also change. For example, major investments in China and the United States are likely to shape research priorities directly in technical fields as well as indirectly influence researchers in social sciences and other disciplines.

Other entities in the stakeholder ecosystem include nonprofit or hybrid organizations like the Partnership on AI or AI Now, some of which have formed in recent years specifically to address AI ethics issues. Even the process of creating ethics documents can provide motivations for collaboration, new partnerships, and new organizations.

## 3.1 Factors that Predict Success

What counts as success depends on the kind of document, the sector of its authors, and its desired goals. However, we can hypothesize a few characteristics worth observing that might suggest that a document is more likely to achieve its purposes:

*3.1.1 Engagement with Law and Governance.* Engagement with law and governance issues is an indication of seriousness. Documents in deep dialogue with national and international laws, standards, and policy strategies may be more likely to translate effectively in the governance landscape than documents that simply articulate broad principles without context or specificity. For this reason, our research considers whether documents reference specific laws and policies rather than merely mention the general importance of complying with law.

*3.1.2 Specificity.* A related issue is specificity, including the level of detail of recommendations and operationalization strategies contained within the document [2]. For example, the UK's [5] national

AI strategy recommends that the "Centre for Data Ethics and Innovation investigate the Open Banking model… for the secure sharing of personal data beyond finance," identifying both an actor and a model to implement towards securing privacy. On the other hand, Intel's public policy-oriented recommendation that "robust privacy regulatory frameworks for the protection of personal data and cybersecurity should also apply to AI implementations" appears less immediately actionable [6]. Of course, a more general principle can be a starting point for further specification by an organization.

*3.1.3 Reach.* The reach and exposure of a document also may correlate with its uptake. Documents that are highly cited or accessed by the public or other stakeholders may be seen as more valuable and of higher quality. Some documents, like the *Montréal Declaration for Responsible AI*, even accrue signatories, another way of measuring reach. A first-mover advantage is potentially important here as well.

*3.1.4 Enforceability and Monitoring.* The degree to which enforceability and monitoring are described also indicate signs of seriousness [8]. An organization willing to subject itself to costly scrutiny, especially independent external scrutiny, is less likely to be engaged in pure signaling. However, there are open questions about the best forms of governance and whether strategies proposed, e.g., by Google, have achieved credibility [10].

*3.1.5 Iteration and Follow-Up.* Finally, plans for iteration and follow-up suggest that an organization is not producing the document for a one-time statement, but intends to carry out the work in the future and make improvements. This could include further operationalizing a preliminary policy framework, pilot testing a strategy, or seeking the ongoing advice of experts or public stakeholders as AI and its social, ethical, and policy implications evolve.

## 4 CONCLUSION

In this paper, we have shared preliminary ethics and policy themes which is part of a study involving more than 80 AI documents produces by governments, private organizations, and NGOs. We reviewed possible challenges associated with the relative homogeneity of the documents' creators and provided a novel typology of motivations that we believe characterize these documents.

We also discussed the varied impacts these documents may have on the global AI ethics and policy governance landscape, including what factors are relevant to assessing whether a given document is likely to be successful in achieving its goals. Going forward, we anticipate that researchers and practitioners will continue to explore these issues as AI's ethical, social, and policy impacts increase, and as AI governance moves from principles to policies.

## ACKNOWLEDGMENTS

## REFERENCES

[1] Carolyn Abbot. 2012. Bridging the Gap - Non-state Actors and the Challenges of Regulating New Technology. *Journal of Law and Society* 39, 3 (Sept. 2012), 329–358. https://doi.org/10.1111/j.1467-6478.2012.00588.x

[2] Corinne Cath, Sandra Wachter, Brent Mittelstadt, Mariarosaria Taddeo, and Luciano Floridi. 2018. Artificial Intelligence and the 'Good Society': the US, EU, and UK approach. *Science and Engineering Ethics* 24, 2 (April 2018), 505–528. https://doi.org/10.1007/s11948-017-9901-7

[3] Future of Life Institute. 2017. Asilomar AI Principles. , 2 pages.

[4] Daniel Greene, Anna Lauren Hoffmann, and Luke Stark. 2019. Better, Nicer, Clearer, Fairer: A Critical Assessment of the Movement for Ethical Artificial Intelligence and Machine Learning. *Critical and Ethical Studies of Digital and Social Media* (2019), 10.

[5] House of Lords, Select Committee on Artificial Intelligence, UK. 2018. *United Kingdom: AI in the UK: Ready, Willing and able?* Technical Report. House of Lords, Select Committee on Artificial Intelligence, London, UK. 183 pages.

[6] Intel. 2017. *Artificial Intelligence: The Public Policy Opportunity.* Technical Report. Intel, Santa Clara, CA. 12 pages.

[7] Anna Jobin, Marcello Ienca, and Effy Vayena. 2019. The global landscape of AI ethics guidelines. *Nature Machine Intelligence* (Sept. 2019), 1–11. https://doi.org/10.1038/s42256-019-0088-2

[8] Ans Kolk and Rob van Tulder. 2002. The Effectiveness of Self-regulation:: Corporate Codes of Conduct and Child Labour. *European Management Journal* 20, 3 (June 2002), 260–271. https://doi.org/10.1016/S0263-2373(02)00043-9

[9] Emma Martinho-Truswell, Hannah Miller, Isak Nti Asare, André Petheram, Richard Stirling, Constanza Gómez Mont, and Cristina Martinez. 2018. *Mexico: Towards an AI Strategy in Mexico: Harnessing the AI Revolution.* Technical Report. British Embassy in Mexico, Oxford Insights, and C Minds, Mexico City, Mexico. 52 pages.

[10] Cade Metz. 2019. Is Ethical A.I. Even Possible? *The New York Times* (March 2019).

[11] Microsoft. 2018. *The Future Computed: Artificial Intelligence and its role in society.* Microsoft, Redmond, WA.

[12] Ministry of Economic Affairs and Employment, Finland. 2017. *Finland: Finland's Age of Artificial Intelligence Turning Finland into a leading country in the application of artificial intelligence.* Technical Report 47/2017. Ministry of Economic Affairs and Employment, Helsinki, Finland. 76 pages.

[13] Ministry of Economic Affairs and Employment, Finland. 2018. *Finland: Work in the age of artificial intelligence: four perspectives on economy, employment, skills and ethics.* Technical Report 21/2018. Ministry of Economic Affairs and Employment, Helsinki, Finland. 60 pages.

[14] Ministry of Finance and Ministry of Industry, Danish Government. 2019. *Denmark: National Strategy for Artificial Intelligence.* Technical Report. Ministry of Finance and Ministry of Industry, Business and Financial Affairs, Copenhagen, Denmark. 74 pages.

[15] National Science and Technology Council, United States. 2016. *United States: The National Artificial Intelligence Research and Development Strategic Plan.* Technical Report. National Science and Technology Council (NSTC), Networking and Information Technology Research and Development Subcommittee, Washington, D.C. 48 pages.

[16] NITI Aayog. 2018. *India: National Strategy for Artificial Intelligence #AIFORALL.* Technical Report. NITI Aayog, Delhi, India. 115 pages.

[17] OECD. 2019. *OECD Principles on Artificial Intelligence.* Technical Report. OECD, Paris, France. 2 pages.

[18] Office of the Prime Minister, Malta. 2019. *Malta: Towards An AI Strategy.* Technical Report. Parliametary Secretariat for Financial Services, Digital Economy and Innovation, Office of the Prime Minister, Valletta, Malta. 48 pages.

[19] State Council of China. 2017. *China's New Generation of Artificial Intelligence Development Plan.* Technical Report 35. State Council of China, Beijing, China. 28 pages.

[20] The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. 2019. *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, First Edition.* Technical Report. IEEE, Piscataway, New Jersey. 294 pages.

[21] Ben Wagner. 2018. Ethics as an Escape from Regulation: From "Ethics-Washing"to Ethics-Shopping? In *Being Profiles: Cogitas Ergo Sum: 10 Years of Profiling the European Citizen,* Emre Bayamlioğlu, Irina Baraliuc, Liisa Janssens, and Mireille Hildebrandt (Eds.). Amsterdam University Press. https://doi.org/10.2307/j.ctvhrd092