

GLOBAL DIGITAL POLICY SNAPSHOT

National AI Strategies & Human Rights

JUNE 2020

INSIDE THIS BRIEF:

- *National AI Strategies and the Global State of Play*
- *Human Rights Implications and Risks*
- *Recommendations for Policymakers*
- *Diving Deeper on Human Rights and AI Governance*

KEY TAKEAWAYS

Challenge: Governments have begun to develop national artificial intelligence strategies to correspond with broad societal adoption of AI technologies, but these strategies have failed to adequately incorporate human rights principles.

State of Play: Several countries have made efforts to acknowledge the relevance of human rights in the context of their AI strategy; however, additional work is needed to ensure that the human rights framework plays a more central role in guiding governance and regulation of AI, as well as the responsible deployment of AI.

Recommendations: As policymakers develop national AI strategies, they should:

- Make explicit reference to human rights and elaborate potential risks;
- Outline and commit to concrete steps to protect human rights;
- Build-in incentives and sector-specific requirements for rights-respecting practice;
- Set out grievance and remediation processes for human rights impacts;
- Lead in regional and international AI policy development, reinforcing international human rights principles; and
- Engage in multistakeholder drafting processes that include human rights experts.

INTRODUCTION

As the deployment of technologies powered by artificial intelligence (AI) accelerates globally, the implications for human rights continue to expand and evolve. While AI is contributing to profound technological progress across an expansive array of disciplines, it also provides governments and non-state actors with ever more powerful capabilities for surveilling, censoring, manipulating, and repressing citizens. Although the risks to human rights are widely acknowledged and well documented, applicable legal and policy safeguards have not materialized at pace.

In the absence of more robust AI policy regimes and rights protections, a growing number of governments have created “national AI strategies,” which reflect key priorities related to the development and deployment of AI, and often guide a country's AI-related research investments. If done well, these documents can provide policymakers, industry, and researchers with guiding principles for responsible development and adoption of AI technologies. Unfortunately, most national AI strategies produced to date have largely failed to meaningfully address normative concerns, let alone incorporate protections for human rights.

National AI strategies present an under-utilized opportunity to advance norms and principles that ensure technological development is ethical and rights-respecting. This issue of the Global Digital Policy Snapshot describes the current landscape of national AI strategies; evaluates the extent to which human rights have been incorporated into them globally; and identifies six principles for making strategies more rights-respecting.

GLOBAL STATE OF PLAY

Since 2017, more than 30 governments have published national AI strategies, and more are currently in development. The common objective of these strategies is to lay out priorities that can shape the country's approach to AI development and deployment. While the strategies vary in scope, a [recent assessment](#) by GDPi and Global Partners Digital noted thirteen common elements, the most prominent of which are attention to research investment; private-sector and governmental applications; ethics; regulation; and worker skills and talent.

In addition to providing guidance to policymakers and governments, national AI strategies offer useful information to nongovernment actors as well. They signal to private-sector stakeholders how future regulations may impact business strategies. They can also equip advocates and civil society organizations with public commitments against which governments can be held accountable.

When it comes to inclusion of and engagement with human rights principles, most national strategies to date are underwhelming. Instead national AI strategies commonly highlight legitimate national policy concerns regarding, for example, economic development or workforce development but neglect or exclude human rights.

In the absence of clear human rights principles, some countries' strategies reference an “ethics” framework, or focus on “human centered” approaches to AI. The problem with these approaches is that they are often very loosely defined, and their foundations are unclear. Notably, while all countries have obligations under international human rights law and conventions, new ethical frameworks are unlikely to find such wide legitimacy and recognition.

While the existing body of national AI strategies seems to recognize the relevance of human rights and normative concerns, efforts to meaningfully protect rights have been largely sidestepped.

HUMAN RIGHTS IMPLICATIONS & RISKS

AI poses multiple risks for human rights, most notably, risks to privacy, freedom of expression, freedom of assembly and association from new forms of AI-based surveillance and censorship.

Algorithms used for decision making can often entrench existing biases, or create new ones. Facial recognition software poses questions not only around racial and ethnic bias and discrimination, but also questions of privacy and autonomy. Automation driven by AI can threaten the right to work. Algorithmic systems also present deep challenges with respect to core democratic concepts of transparency and accountability in governance. On the other hand, if well-governed, AI may create opportunities in education, health care, and other sectors that bolster people's rights, health, wealth, and quality of life. As national AI strategies emerge as policy instruments, it is important that these documents account for the essential role of human rights norms in guiding the present and future of digital society.

While the majority of existing national AI strategies mention human rights, very few engage deeply with how rights should be protected as regulations take shape. This is generally true even among countries strongly committed to human rights. Of the 32 existing strategies worldwide only 20 explicitly mention the existing human rights framework, and of these no more than half a dozen take meaningful steps to ensure that human rights are protected. Much more commonly, human rights are mentioned in passing, such as in the introduction or as a general reminder of the country's commitment. In other cases, specific rights, such as the right to privacy, are discussed in some detail while other rights are not mentioned at all. Additionally, even when rights are mentioned, there is rarely a clear, detailed plan for how to protect them. These deficiencies can produce sobering consequences in the long term, particularly as the world has seen a growing willingness among illiberal governments to harness technological power to suppress free expression, free association, and political participation.

Simply put, significant steps can and should be taken to integrate human rights much more meaningfully. A more robust approach would acknowledge implications for human rights in every applicable section or chapter of the national strategy, with dedicated reference to the human rights framework, concrete actions the government will take to protect rights, metrics for evaluation, and incentives or requirements for adherence. Moreover, strategies should highlight existing mechanisms for remedy and other government processes or consequences that may be implicated.

In a few noteworthy cases, several countries have already made substantial efforts to integrate human rights more deeply into their national AI strategies. The Netherlands, for example, dedicates an entire section to the potential risks and opportunities of AI for human rights. They also have a detailed list of actions the government is taking to try to protect these rights, such as investments in research on AI's impact. Meanwhile the Norwegian strategy outlines seven standards for AI, grounded in human rights principles, including that AI solutions must respect human autonomy and control and facilitate inclusion, diversity and equal treatment. Norway's strategy also features commitments about what the government will do to ensure that AI is trustworthy and rights-respecting. In other cases, the countries' core strategy documents do not include such strong engagement with human rights, but subsequent guidance or implementation documents have helped elevate human rights principles with growing awareness of their importance in the context of AI.

These examples illustrate the potential for national strategies to meaningfully incorporate human rights principles, and the ways that the human rights framework can serve both as a basis for assessing societal impacts of AI and as a guide for regulating its deployment. It remains clear, however, that this potential has not been fully realized. With a more intentional approach, national AI strategies can play a greater role in advancing a rights-respecting future.

RECOMMENDATIONS FOR POLICY MAKERS

Policymakers should follow these six recommendations in their national AI strategies:

1. **Make explicit reference to human rights and elaborate potential risks:** Include human rights explicitly and throughout the strategy. Specific human rights risks should be elaborated, including potential impacts on at-risk, vulnerable and marginalized communities, along with a commitment to mitigate those risks.
2. **Outline and commit to concrete steps to protect human rights:** National strategies should include specific goals, commitments or actions that will ensure that human rights are protected.
3. **Build-in incentives and sector-specific requirements for rights-respecting practice:** Governments should incentivize human rights-respecting practices across all sectors, and require concrete steps in specific sector practices to ensure protection of human rights.
4. **Set out grievance and remediation processes for human rights impacts:** Governments should assess existing grievance and remedial processes and determine whether those processes need revision in light of the particular nature of AI, so that complaints concerning AI can be received and remediated.
5. **Lead in regional and international AI policy development:** Governments should proactively engage in regional and global fora and processes relating to AI governance, and promote human rights-respecting approaches and outcomes.
6. **Engage in multistakeholder drafting processes that include human rights experts:** When drafting a national strategy, the government should engage with cross-disciplinary, cross-sector experts, including human rights experts.

National AI Strategies and Human Rights: A Review, GDPi and Global Partners Digital.

Governing Artificial Intelligence: Upholding Human Rights and Dignity, Data & Society.

Human Rights in the Age of Artificial Intelligence, AccessNow.

Artificial Intelligence and Human Rights, Eileen Donahoe & Megan Metzger.

DIVING DEEPER ON HUMAN RIGHTS & AI GOVERNANCE

AUTHORS

Eileen Donahoe is Executive Director of the Global Digital Policy Incubator. She previously served as the U.S. Ambassador to the UN Human Rights Commission, and serves in leadership and advisory roles across the global democracy and human rights community.

Megan Metzger is a Research Scholar and Associate Director of Research for the Global Digital Policy Incubator. She completed a PhD in Politics at NYU as a member of the Social Media and Political Participation Lab and her research focuses on how technology impacts rights and political behavior.

Kip Wainscott is a Senior Advisor for the Global Digital Policy Incubator. A lawyer and policy professional with experience in government, civil society, and the private sector, he has worked extensively on issues concerning technology's impact on democracy and human rights.

This GDPi Policy Snapshot is made possible by the generous support of the International Center for Not-for-Profit Law (ICNL). The contents are the responsibility of GDPi and do not necessarily reflect the views of ICNL.

ICNL
INTERNATIONAL CENTER
FOR NOT-FOR-PROFIT LAW