THE STATE OF AI POLICY: THE DEMOCRATIC VALUES PERSPECTIVE

In a world where data means power, vast amounts of data are collected every day by both private companies and government agencies, which then use this data to fuel complex systems for automated decision-making now broadly described as "Artificial Intelligence." Activities managed with these AI systems range from policing to military, to access to public services and resources such as benefits, education, and employment. The expected benefits from having national talent, capacity, and capabilities to develop and deploy these systems also drive a lot of national governments to prioritize AI and digital policies. A crucial question for policymakers is how to reap the benefits while reducing the negative impacts of these sociotechnical systems on society.

Merve Hickok* & Marc Rotenberg**

TPQ
Winter 2021/22

* Merve Hickok is CAIDP Research Director and Founder of the AIethicist.org.
** Marc Rotenberg is Founder and President of CAIDP and a faculty member at Georgetown Law.

The past year has produced rapid changes in the world of AI policies and practices. National governments and international organizations are moving to create new frameworks to tip the balance in favor of benefits of artificial intelligence. There is consensus that wider adoption of these systems requires a level of trust from public, which in turn requires human-centric systems. But "trustworthy" and "human-centric" cannot simply be slogans. For a system to be trustworthy, it must be subject to independent evaluation. It should produce results that can be verified, replicated, and proven. "Human-centric" necessarily requires the centrality of human agency. Unaccountable, autonomous devices making decisions about people stand at the opposite end of the spectrum from human-centric, trustworthy AI.

From a civil society perspective, so far, the lack of regulations which would ensure oversight over these systems, coupled with minimum transparency have been an increasing source of concern to those working to keep governments, public and private entities accountable. With the expanding adoption of AI systems, the implications over fundamental human rights, democratic values and rule of law are at the core of the discussions. The right to privacy and other rights, including the rights to health, education, due process, freedom of movement, freedom of peaceful assembly and association, and freedom of expression can all be impacted by the way AI systems are deployed.

The Center for AI and Digital Policy (CAIDP) has set forth an objective in 2020 to independently assess global progress toward human-centric and trustworthy AI. Researchers at CAIDP agreed that such progress required an extensive empirical survey of AI policies in countries across the world. The survey also required metrics which corresponded to widely accepted frameworks and principles, allowing for quantification of commitments and practices.[1] The result, AI and Democratic Values Index, has these objectives: (1) to document the AI policies and practices of influential countries, based on publicly available sources, (2) to establish a methodology for the evaluation of AI policies and practices, based on global norms, (3) to assess AI policies and practices based on this methodology and to provide a basis for comparative evaluation, (4) to provide the basis for future evaluations, and (5) to ultimately encourage all countries to make real the promise of AI that is trustworthy, human-centric, and provides broad social benefit to all.

The AI Index is based on such well-established norms and sources of authoritative assessments. The OECD/G20 AI Principles[2] provided a starting point, as did

[1] Marc Rotenberg, "Time to Assess National AI Policies and Practices"*, Communications of the ACM, 24 November 2020,* https://cacm.acm.org/blogs/blog-cacm/248921-time-to-assess-national-ai-policies/fulltext

[2] OECD AI Principles, (2019). https://oecd.ai/en/ai-principles

the Universal Declaration for Human Rights[3], the most widely recognized legal instrument for fundamental rights. We developed a methodology, drawing on the work of international human rights organizations and data protection experts. We revised questions as our work progressed, and new factors were uncovered. We recognized early on the difference between a country's endorsement of a key principle, such as "fairness," and a country's implementation of that principle. Endorsement is easy to measure, implementation, not so much. In highlighting this distinction, we hope others will also look more closely at the difference between what countries say and what they do, all with the larger purpose of closing that gap. And we knew we could not look at the practices of all countries, so we chose those countries (again relying on objective metrics) that we thought would be most impactful.[4] In 2021, CAIDP published its analysis of 30 countries against 12 metrics. In 2022, the analysis has been expanded to 50 countries with the assistance of more than 100 policy experts. Our research team reviewed national AI strategies, commitments to international policy frameworks, adherence to democratic norms such as data protection and transparency, statements from policy makers, and the activities of government agencies. The country report narratives provided the basis for quantitative assessment, followed by ratings and rankings for individual countries.

"How a national government handles the questions of accountability, power and progress is directly relevant to how much it respects democratic values and fundamental rights."

Some AI systems used by public or private actors carry huge risks for society and democratic values if their development and deployment do not center the dignity and rights of natural persons. How a national government handles the questions of accountability, power and progress is directly relevant to how much it respects democratic values and fundamental rights. Of course, no country operates in a vacuum. International organizations, such as the OECD, the G7 and G20, UNESCO, the European Union, and the Council of Europe can establish policy frameworks and legal standards that promote beneficial uses of technologies and curb the dangerous ambitions of some governments. This is particularly evident in the AI policy field where emerging global norms could, if effectively implemented, prevent the use of AI techniques for social scoring, mass surveillance, and autonomous weapons. Therefore, a look at the state of global AI policy also requires an understanding of progress in these organizations.

[3] United Nations, *Universal Declaration of Human Rights.*

[4] CAIDP, *Artificial Intelligence and Democratic Values Index 2020 (2020).* https://www.caidp.org/reports/aidv-2020

From an individual country perspective, this year's edition of AI & Democratic Values Index[5] places Canada, Germany, Italy, and Korea in the top tier for their global leadership on AI policy, their commitment to democratic values, and meaningful engagement with the public on proposed AI strategies. Other important factors for top rankings were a well-established data protection infrastructure, support for algorithmic transparency, and a commitment to fairness, accountability, and transparency for AI systems. In the past year, Canadian authorities determined that ClearviewAI was a form of mass surveillance and violated the privacy and data protection rights of Canadians. Germany continued its leadership on AI policy in the European Union, emphasizing protection for fundamental rights and ongoing public participation on AI policy development. As host of the G20 summit, Italy advanced AI policy proposals, emphasizing data protection and gender equality, diversity, and inclusion. Korea introduced new requirements for AI impact assessments, published guidance on Personal Information Protection, and expanded algorithmic transparency.

China has also adopted sweeping new laws for both data protection and the regulation of recommendation algorithms. Although the privacy rules look very similar to the GDPR and the regulation for the governance of recommendation algorithms share similar ambitions to those proposals pending in both the European Union and the U.S Congress, there are real concerns about AI policies that are intended to favor the government in power. In that context, the goals of transparency and accountability is offset by the inherent bias of such a legal structure.

Despite its private sector's global edge in AI innovation, the U.S. ranked in third tier in the Index. The absence of a legal framework to implement AI safeguards and a federal agency to safeguard privacy raises concerns about the ability of the U.S. to monitor AI practices.

On the global front, the UNESCO Recommendation on the Ethics of AI, adopted by 193 countries in November 2021, was the single most significant AI policy development of the past year.[6] The UNESCO Recommendation speaks directly to the widespread – and widely shared – aspiration of countries that AI should benefit humanity. In the AI policy field that barely existed a few years ago, the UNESCO AI Recommendation is a remarkably comprehensive AI policy framework, touching upon established concerns regarding AI systems, such as fairness, accuracy, and transparency, and emerging issues, including gender equity and sustainable development. UNESCO's proposal for Ethical Impact Assessment provides a

[5] CAIDP, *Artificial Intelligence and Democratic Values Index 2021 (2021).* https://www.caidp.org/reports/aidv-2021

[6] UNESCO, *Recommendation on the ethics of artificial intelligence* (2021), https://en.unesco.org/artificial-intelligence/ethics

powerful new tool to assess, in advance, the consequences of the deployment of AI systems. Recognizing the importance of the first global framework for AI ethics, CAIDP has this year altered one of the metrics to take account of the significance of the UNESCO Recommendation on AI. It is a development worth acknowledging and celebrating. In future reports, we will likely add another metric to assess the far more challenging issue of implementation.

"As the field of AI policy rapidly matures, we observe the growing presence of judicial decisions, now shaping the laws of algorithms. In several cases, including the secretive evaluation of employee performance, courts have rejected opaque automated decisions. These judgements are based on well-established legal frameworks, such as the GDPR, though we see also legislative efforts to make automated decision-making with AI techniques more accountable."

Since the publication of our last report, we also note the introduction of the European Commission proposal for the regulation AI. The Commission has set out a comprehensive, risk-based approach that could extend the "Brussels Effect" to the global governance of AI. The European Parliament has also signaled its intention to strengthen key provisions, and likely will prohibit the use of AI techniques for remote biometric identification. Meanwhile, the Council of the European Union, under the Presidency of Slovenia and now France, have proposed additional texts that would among other changes, extend the prohibition on social scoring to private companies as well as public agencies.

2021 also marked the adoption of Resolution 473 in Africa, concerning the need to undertake a study on human and peoples' rights and artificial intelligence. The African Commission on Human and Peoples' Rights called on State Parties "to ensure that the development and use of AI, robotics and other new and emerging technologies is compatible with the rights and duties in the African Charter and other regional and international human rights instruments, in order to uphold human dignity, privacy, equality, non-discrimination, inclusion, diversity, safety, fairness, transparency, accountability and economic development as underlying principles

that guide the development and use of AI, robotics and other new and emerging technologies."[7]

The Council of Europe (COE), continent's leading human rights organization[8], comprised of 47 member states, had established the Ad Hoc Committee on Artificial Intelligence (CAHAI) in September 2019.[9] The mandate of the CAHAI was to "examine the feasibility and potential elements on the basis of broad multi-stakeholder consultations, of a legal framework for the development, design and application of artificial intelligence, based on the Council of Europe's standards on human rights, democracy and the rule of law."[10] The CAHAI held its final meeting in December 2021.[11] At the end of the meeting, the CAHAI adopted the "Possible elements of a legal framework on artificial intelligence, based on the Council of Europe's standards on human rights, democracy and the rule of law." The CAHAI framework contains an outline of the legal and other elements which in the view of the Committee could be included in legally binding or non-legally binding instruments of COE that will make up an appropriate legal framework (possibly a convention) on AI. The CAHAI framework is now submitted to the Committee of Ministers for further consideration. The Committee has already emphasized in 2021 that such systems should be developed and implemented in accordance with the principles of legal certainty, legality, data quality, non-discrimination, and transparency.[12]

We also noted the growing conflict over the deployment of facial recognition for mass surveillance. While the European Parliament voted to ban the use of AI technology for this purpose, many governments and private companies pushed forward new systems for surveillance in residential communities, inside school classrooms, and at public parks. These are not the CCTV cameras of old, but sophisticated image processing systems, designed specifically to identify individuals in public spaces by name. In some countries, this system of unique identification is then tied to elaborate government databases for scoring people based on their allegiance to

[7] African Commission on Human and Peoples' Rights, *473 Resolution on the need to undertake a Study on human and peoples' rights and artificial intelligence (AI), robotics and other new and emerging technologies in Africa - ACHPR/ Res. 473 (EXT.OS/ XXXI)* (25 February 2021).

[8] Council of Europe, "Who we are," https://www.coe.int/en/web/about-us/who-we-are

[9] Council of Europe, "The Council of Europe established an Ad Hoc Committee on Artificial Intelligence – CAHAI," (11 September 2019), https://www.coe.int/en/web/artificial-intelligence/-/the-council-of-europe-established-an-ad-hoc-committee-on-artificial-intelligence-cahai

[10] Council of Europe, "CAHAI - Ad hoc Committee on Artificial Intelligence," https://www.coe.int/en/web/artificial-intelligence/cahai

[11] Council of Europe, "The CAHAI held its 6th and final plenary meeting," (2 December 2021), https://www.coe.int/en/web/artificial-intelligence/-/outcome-of-cahai-s-6th-plenary-meeting

[12] Council of Europe, "Declaration by the Committee of Ministers: the use of computer-assisted or AI-enabled decision making by public authorities in the area of social services must respect human rights" (17 March 2021), https://www.coe.int/en/web/artificial-intelligence/newsroom/-/asset_publisher/csARLoSVrbAH/content/declaration-by-the-committee-of-ministers-the-use-of-computer-assisted-or-ai-enabled-decision-making-by-public-authorities-in-the-area-of-social-servi

the government in power. It is a form of social control beyond the imagination of even Jeremy Bentham's panoptic prison design.

We called attention also to the unfortunate failure of negotiators at the UN conference in late December to make progress on a proposal to limit, or better to prohibit, the use of lethal autonomous weapons. This occurred in the same year that the United Nations suggested the first use of autonomous drone swarms to target and kill retreating military forces in the civil war in Libya.

As the field of AI policy rapidly matures, we observe the growing presence of judicial decisions, now shaping the laws of algorithms. In several cases, including the secretive evaluation of employee performance, courts have rejected opaque automated decisions. These judgements are based on well-established legal frameworks, such as the GDPR, though we see also legislative efforts to make automated decision-making with AI techniques more accountable. We report these outcomes favorably as algorithmic transparency remains one of our key metrics for the evaluation of AI policies and practices.

In addressing the need to advance democratic values in the age of AI, the ability of the European Union, the United States, and allies to work in common purpose remains central. On that front, the past year provides reason for both optimism and concern. The EU and the U.S. launched a Trade and Technology Council in 2021 that set out a common framework on AI policy that could promote further transatlantic cooperation.[13] The good news is that "human rights" and "democratic values" undergird many of the proposals. Top officials in the Biden Administration also expressed support for the EU AI Act, a key legislative framework that will likely move forward in 2022.

At the same time, the future of the EU AI Act is not certain, as some politicians have made the mistake of assuming it is possible to trade the protection of rights for innovation. Technologies that fail to protect rights are not innovative, they are oppressive and stifling. On the U.S. side, several federal agencies have initiated AI-related "listening sessions," but the necessary work of establishing legal standards to protect democratic values has yet to begin.

Still, our survey of national AI policies and practices also revealed the hard work of many NGOs, advocates, academics, and government officials, around the world, who have fully engaged the challenges that AI poses and are prepared to stand on

[13] European Commission, *EU-US Trade and Technology Council: Commission launches consultation platform for stakeholder's involvement to shape transatlantic cooperation* (October 2021). https://ec.europa.eu/commission/presscorner/detail/en/IP_21_5308

the front lines in defense of fundamental rights. The remarkable progress made by the "ReclaimYourFace" campaign in Europe, and similar campaigns in Africa, Asia, and Latin America speak to a rapidly growing public recognition that not all technologically transformative impacts should be welcome.

This point was made clear during the past year with the call from Michelle Bachelet, United Nations High Commissioner on Human Rights, for a prohibition on AI techniques that fail to comply with international human rights law. Commissioner Bachelet stated, "The higher the risk for human rights, the stricter the legal requirements for the use of AI technology should be."[14]

There is a growing understanding that "red lines" are necessary to safeguard fundamental rights. And in that recognition may be found also the key to aligning AI policies and practices, to narrowing the gap between the world of AI as it is and the world of AI we wish to inhabit. If AI is to remain human-centric, then we must determine the appropriate applications of AI.

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