

CAIDP Release of Report – *AI and Democratic Values*

By Afi C. Blackshear

On Monday, February 21, 2022, the Center for AI and Digital Policy marked the release of the report *Artificial Intelligence and Democratic Values Index* with a panel of experts, and the Vice President of the European Parliament **Eva Kaili** covering the current state of AI policy. The event, moderated by CAIDP Program Director **Karine Caunes**, opened with remarks from CAIDP Research Director **Merve Hickok**.

Hickok summarized key findings of the 2021 report, which expanded its scope of coverage from 30 to 50 countries in an effort to measure how effectively countries safeguard human rights and democratic principles. Canada, Germany, Italy, and South Korea scored in the top tier for their commitment to democratic values, as well as their meaningful engagement with the public on proposed AI policies. They also established data protection infrastructure, supported algorithmic transparency, and committed to fairness and accountability mechanisms for AI systems.

*“The Index found global support for strong frameworks to regulate AI. At the same time, the threats to fundamental rights, enabled by AI techniques, are on the rise.” –
CAIDP Research Director Merve Hickok*

Hickok also discussed how AI Superpowers — the United States and China — made progress, but continued to score poorly, landing in tiers 3 and 4, respectively. The United States’ rating improved as the Biden Administration stressed democratic values, public participation, and algorithmic transparency, but actual policymaking was “opaque.” China continued the widespread use of AI techniques for facial recognition against ethnic minorities and political protestors, as well as to socially score people. Still, China’s ranking improved because of its endorsement of UNESCO Recommendations on AI Ethics, but Hickok asserted that “an urgent need to assess China’s actual practices against these global standards” remains.

The event then turned to keynote speaker Eva Kaili, Vice President of the European Parliament. Kaili explained how AI’s exponential proliferation represents “revolutionary” opportunity for Europe and the dissemination of democratic ideals of like-minded nations, especially when considering the \$15 trillion AI technologies are estimated to add to the global economy. Kaili described the European Union’s AI Act that is currently under negotiation. She called it a “far-reaching effort to have a trustworthy AI that puts people in the center of this technological revolution.” She called the AI Act a unifying framework in contrast to the hundreds of AI ethics frameworks, arguing for society’s need to agree on one basic and binding legal framework for all to follow. In her view, the AI Act is a positive step toward that end, and she expressed hope that the AI Act itself will mirror the impact of the European Union’s General Data Protection Regulation (GDPR) and establish a baseline for other nations to follow.

Kaili supported the need for humans to remain in the loop of AI systems, and the importance of embedding values, such as human oversight, transparency, and accountability. Absent these standards, she noted how nations risks “embarrassing” themselves like the United Kingdom has with its implementation of a grading system that failed, incorrectly assigning grades to students. The AI Act, thus, crystallizes principles into a horizontal binding framework to ensure appropriate safeguards that regulate applications that can be harmful in the European market.

“High risk systems should be subject to strict standards for training, for validation and testing, for risk-management measures, and for disparities in oversight measures.” – Vice President of the European Parliament Eva Kaili

Kaili then discussed the benefits of adopting a risk-based approach, wherein high risk uses ought to be subject to “strict standard requirements for training, for validation and testing, for risk-management measures, and for disparities in oversight measures.” However, while the principles to be applied to high-risk uses are clear, Kaili elaborated how there is currently little consensus for what constitutes a high-risk use. Regardless, she believes that subliminal manipulation, social scoring by public authorities, real-time remote biometric identification in public spaces, the management and operation of critical infrastructure, education, employment and recruitment, access to public and private services, law enforcement, predictive AI, predictive policy, and the administration of justice should be considered high risk. “We’ve seen, already, negative examples that we need to avoid.”

On the other hand, she described the more common “minimal risk” applications of artificial intelligence, such video games, chat boxes, or spam filters, which Kaili asserted should also be subject to disclosure, cybersecurity, and transparency standards. “However, the main debate in Europe is what counts as a risky AI system,” which, she concluded, is why it is important for all people to participate in policy formulation.

*“The main debate in Europe is what counts as a risky AI system.”
– Vice President of the European Parliament Eva Kaili*

Kaili finished her speech by addressing the need for collaboration and consistent methodologies to better assess the implementation of various promulgated principles and standards. “The patchwork of initiatives is coming together as pieces of this puzzle of what we’re trying to accomplish in Europe to have this big picture of what we are trying to achieve and update. I think, toward these goals, the *AI and Democratic Values Index* is really an important step.”

Karine Caunes then transitioned the event to the panel discussion. Panelists included Fanny Hidvéghi, Manager of Access Now Europe Policy; Eduardo Bertoni, a representative of South

America at the Inter American Institute of Human Rights and former director of the Argentina’s Data Protection Authority; Jibu Elias, head of National AI Portal of India; Stuart Russell, Professor of Computer Science of University of California, Berkeley and Director of the Center for Human-Compatible AI; and Merve Hickock, the Center’s Research Director.

*“The challenge is how are we going to use artificially intelligent technologies in a way that will not damage human rights.” –
Eduardo Bertoni*

Building on Kaili’s keynote, the panelists exchanged perspectives on the current state of AI policy worldwide. Eduardo Bertoni commenced the exchange, describing key challenges ahead for policymakers who wish to uphold and promote democratic values in AI in Latin America. “In Latin America – and this is global – Artificial Intelligence is growing not only in the private sector, but also in the public sector.” From his view, this growth can be a tool to either damage or enhance, protect, and strengthen human rights. “The challenge is how are we going to use artificially intelligent technologies in a way that will not damage human rights.” Bertoni agreed with Kaili regarding the importance of creating consistent standards. “That’s why I think the report is very useful, because it establishes a methodology that countries can use to see how they are doing in respecting human rights.”

*“If you cannot measure it, you cannot improve it.”
- Jibu Elias*

According to Jibu Elias, the biggest challenges for policymakers are the speed of technology and the speed of policymaking. “We live in a world where it’s all about policy planning, but we need to shift into real-time policymaking.” Elias stressed the difficulty “radically” unrepresentative data has in undercutting effective policymaking, noting the digital divide that results in the average household in the United States being able to produce six data points per second compared to lower income countries that, on average, produce zero. “If you cannot measure it, you can’t improve it.”

*“We have long advocated for ‘red lines’ for AI.”
- Fanny Hidvegi*

Fanny Hidvéghi spoke next, responding, “Absolutely yes,” to Karine Caunes’s question on the need for “red lines” for high risk uses of AI. “I believe that one of the biggest opportunities the EU AI Act brings is prohibitions, which is something we have long been advocating for. But the

current proposals definitely do not go far enough.” Hidvéghi expressed disappointment in law enforcement and national security exemptions that “sometimes make red lines moot.”

The panel turned to Professor Stuart Russell, who affirmed the need for bans, particularly on the use of AI to impersonate human beings. He expressed support for the “fundamental human right” to know whether a person is communicating with a human being or interacting with AI. “Without that,” Russell asserted, “the whole concept of human dignity itself is likely to be invalidated.”

Recalling a conversation on the metaverse, Russell described the “catastrophic” risk reinforcement learning algorithms can have on democracy. According to Russell, these algorithms learn and nudge people to change who they are and push people to become more predictable and, inevitably, more extreme to maximize clicks and engagement. “Many people are living in an alternate universe. . . What [reinforcement learning algorithms] have learned to do is to turn people into extremists of any kind — they do not care what kind of extremist you are — because the more extreme you are, the more likely people are to click on the content that is sent.” Russell argued that these applications are critical elements of high-risk uses of AI discussed in Kaili’s keynote. “Information systems are high risk systems because they can damage your mental integrity.” He also echoed Hidvéghi’s concerns about exemptions, encouraging the EU to stick to its commitment to fundamental rights, “despite many attempts to water down and have carve outs for entire AI systems.”

“Information systems are high-risked systems because they can damage your mental integrity.”
- Stuart Russell

Russell finished his remarks by providing an overview of debates related to lethal autonomous weapons, which he argued are positioned to usher in a new era of mass destruction, wherein the United States and Russia oppose red line prohibitions. He encouraged the public to increase the pressure on policymakers. “Just ask people: do you like the idea of robots hunting and killing human beings? Why aren’t politicians protecting us? I don’t have a good answer, and I don’t think the politicians have a good one, either.”

“Governments should aim high and establish red lines for mass surveillance, for autonomous weapons, and for systems that rank and score citizens”
- Merve Hickok

Merve Hickok concluded the discussion, calling on policymakers to “aim high” as they propose and implement human-centric policy. She, too, affirmed the need for red lines for AI

applications that rank and score citizens, lethal autonomous weapons, and mass surveillance — all of which “converge in providing the means to oppress, exploit, discriminate, or manipulate humans.”

CAIDP Director Marc Rotenberg delivered concluding remarks. Echoing Jibu Elias, Marc explained how, “We know from our background in the human rights world that governments are quite willing to endorse a declaration, but the hard work begins trying to assess implementation and whether progress is being made toward the goals that governments have established. And that is at the heart of this report.”

“General purpose AI systems require the highest degree of regulation as these are the systems that will have the greatest impacts on our society.”
- Marc Rotenberg

Referring to the EU AI Act, Rotenberg said it is the general purpose AI systems that require the highest degree of regulation as “these are the systems that will have the greatest impacts on our society.”

