

CAIDP Update 6.08

February 26, 2024

Welcome to the CAIDP Update. Join us as we explore the constantly changing landscape of AI governance and regulation.

In far East, Japan has established an AI Safety Institute. In the European Union, a centralized European AI Office has been launched. The OECD has initiated an expert group to harmonize AI, data, and privacy policies globally. The United Nations advocates for a Global Digital Compact aimed at fostering an inclusive digital future amidst geopolitical tensions.

In the Middle East, the UAE-UNDP partnership aims to leverage AI for sustainable development goals. Meanwhile, Hong Kong is introducing data privacy checks to safeguard personal information. At CAIDP, we critiqued the effectiveness of the Tech Accord aimed at mitigating AI-generated misinformation in the 2024 elections.

With the Council of Europe AI Treaty approaching the finish line, concerns are growing that private sector AI systems, the most impactful in the world, may be excluded from the global treaty. Following efforts in Europe and the United States, attention is turning to Canada where NGO leaders, AI experts, and human rights experts are urging the Canadian government to back a robust treaty that does not exclude AI systems operated by Big Tech. Below is an opportunity for all to join this campaign.

**Open Letter Urges Canadian
Government to Reject Private
Sector Carve-Out in Council of
Europe AI Convention Negotiations**

BECOME A SIGNATORY

CALL FOR ACTION: Civil Society Leaders Launch Campaign, Urge Canada to Back Strong Global AI Treaty, Oppose Carveouts for Private Sector AI Systems

Civil society leaders in Canada are urging support for a public statement to ensure that the government of Canada commits to safeguarding human rights from AI risks and harms no matter the source. Negotiations are underway at the Council of Europe to finalize the first global treaty on AI. The Council of Europe is an international organization mandated to uphold human rights, democracy and the rule of law.

"Given the profound impacts of AI on human rights and democratic values, any other position of the Canadian delegation would be a violation of the Canadian Charter of Rights and Freedoms," said the organizers. The groups warned, "A hollowed-out Convention will provide little meaningful protections to individuals who are increasingly subject to powerful AI systems prone to bias, human manipulation, and the destabilization of democratic institutions."

References

Tug of war continues on international AI treaty as text gets softened further
Luca Bertuzzi, Euractiv, Feb. 15, 2024

EU prepares to push back on private sector carve-out from international AI treaty,
Luca Bertuzzi, Euractiv, Jan. 12, 2024

EU's AI ambitions at risk as US pushes to water down international treaty
Luca Bertuzzi, Euractiv, June 8, 2023

CAIDP Statements on the Council of Europe AI Treaty

Principles for responsible, trustworthy and privacy-protective generative AI technologies. December 3, 2023. Office of the Privacy Commissioner of Canada

AI POLICY NEWS

Japan Launches AI Safety Institute

Japan's Ministry of Economy, Trade, and Industry (METI) has announced the establishment of the AI Safety Institute, a new entity dedicated to studying and



evaluating the safety of AI. This institute is set to operate under the Information-technology Promotion Agency (IPA), an independent administrative agency, with support from the Cabinet Office, various ministries, agencies, and related organizations.

The creation of the AI Safety Institute reflects Japan's response to the increasing global focus on AI safety. It aims to develop methods for AI safety evaluation and to collaborate with similar institutions globally, including those in the United States and the United Kingdom, to advance standards and evaluation methods for AI safety.

The institute's main activities will include research related to safety evaluations, developing standards, considering implementation methods for safety assessments, and fostering international cooperation with related organizations abroad. This initiative will leverage Japan's domestic and international networks, and expertise in AI, facilitated by contributions from multiple related ministries, agencies, and organizations within Japan.

European Commission Launches AI Office

The European Commission has established the European AI Office as the central hub of AI expertise across the EU. This office is designed to be at the forefront of implementing the AI Act, particularly focusing on general-purpose AI, and aims to "foster the development and use of trustworthy AI, and international cooperation." It positions itself as the foundation for a unified European AI governance system, emphasizing the need for AI to be safe, trustworthy, and beneficial across the 27 Member States.



The European AI Office will have a crucial role in ensuring the AI Act's effective implementation by supporting Member States' governance bodies. It is entrusted with enforcing rules for general-purpose AI models and promoting a culture of trustworthy AI within the EU. The office's duties include "supporting the AI Act and enforcing general-purpose AI rules," thereby contributing to the coherent application of the AI Act across Member States and developing tools for evaluating AI models.

Furthermore, the office is committed to strengthening the development and use of trustworthy AI by "advancing actions and policies to reap the societal and economic benefits of AI across the EU." This involves providing advice on best practices, facilitating access to AI sandboxes for real-world testing, and encouraging the

OECD Forms Expert Group to Harmonize AI, Data, and Privacy Policies



The OECD has announced the formation of an expert group focused on policy synergies in AI, data, and privacy. Recognizing the challenges and opportunities presented by AI's expansion across global sectors, the group aims to foster international multi-stakeholder cooperation in data governance and privacy. The group's core premise is that "data protection laws and privacy considerations have a role to play in building trustworthy AI," underscoring the potential for harmonious development that respects both innovation and individual rights.

CAIDP Founder Marc Rotenberg was named to the Steering Group for the expert panel, along with Yordanka Ivanova (EU DG Connect), Kari Laumann (Norwegian DPA), and Winston Maxwell (Telecom Paris). The co-chairs of the committee are Denise Wong (IMDA/PDPC Singapore), Dr. Clara Neppel (IEEE), and Reuven Eidelman (Israel DPA).

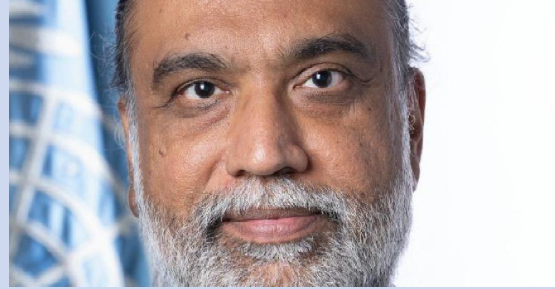
The expert group seeks to bridge the gaps between AI and privacy communities, facilitating a shared understanding and collaborative efforts towards compatible regulations. This initiative reflects a broader recognition of the need for integrated approaches, as "Principles and concepts in data privacy will also need to be mapped to AI principles to ensure alignment and to develop a common language between the communities." With increasing enforcement actions around privacy in the AI sector, the group emphasizes the constructive role of privacy regulations in enabling responsible AI innovation, rather than hindering it.

The OECD has long been a pivotal forum for setting international guidelines on emerging technologies. Notably, in 1980, it established the influential Privacy Guidelines. In a similar vein, the OECD AI Principles of 2019 have garnered endorsement from over 50 countries, underscoring their global impact.

UN's Gill Calls for Global Tech Cooperation Amidst Tensions

Amidst the current geopolitical tensions surrounding technology, UN Secretary-

General's Envoy on Technology Amandeep Singh Gill, in an op-ed, asserts the importance of international cooperation in the realm of tech policy. Despite the intuitive conflict between tech competition and collaboration, Gill states,



"it seems counter-intuitive to speak about building international cooperation on tech policy." This highlights the UN's commitment to leveraging digital cooperation to fulfill its mandates in peace, security, human rights, and sustainable development through the Office of the Secretary-General's Envoy on Technology (OSET), established in 2021.

Looking forward to the Summit of the Future in September 2024, Gill emphasizes the critical need for a "Global Digital Compact" to ensure an "open, free, secure, and inclusive digital future." He underlines the urgency of this mission, pointing out that "2.6 billion people still cannot access the Internet," showcasing the vast digital divide that continues to persist. Furthermore, Gill sheds light on the governance of AI, sharing that "An Advisory Body of independent experts...is working alongside to craft recommendations on the international governance of AI."

UAE and UNDP Partner to Boost AI Development

The UAE's office responsible for AI, Digital Economy, and Remote Work Applications has signed an agreement

with the United Nations Development Programme (UNDP) aimed at enhancing the research and development of AI technologies. This partnership is designed to empower talent and foster innovation in the Arab region and beyond, with a strong focus on achieving sustainable development goals (SDGs).



The collaboration seeks to explore the ethical and responsible use of AI, integrating it across various sectors to make impactful progress that benefits communities globally. It aims to increase the adoption of AI technologies to enhance societal well-being, support sustainable development, and devise innovative solutions to challenges.

Furthermore, the agreement is set to promote AI as a tool for development priority, focusing on exchanging knowledge, supporting digital initiatives, and enhancing decision-making processes with AI technologies. Through this initiative, both organizations aim to leverage AI to accelerate progress toward a sustainable and

inclusive future.

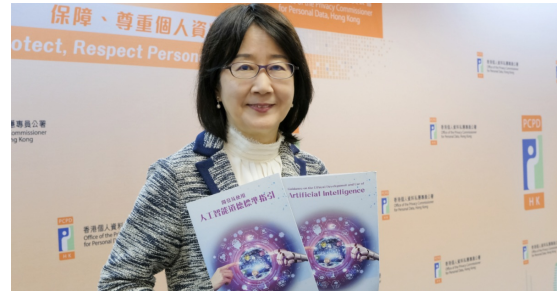
Hong Kong's Privacy Commissioner Conducts AI Compliance Checks

Privacy Commissioner for Personal Data (PCPD) in Hong Kong has completed a comprehensive review of 28

organizations to evaluate their practices in handling personal data within AI systems and their AI governance frameworks. This survey spanned diverse sectors including telecommunications, finance, beauty services, retail, transportation, education, and government.

The PCPD's findings highlighted that out of 21 organizations integrating AI into their operations, 19 had implemented internal AI governance structures, and 10 actively collected personal data through AI services. Specifically, among these 10 organizations, all had adopted stringent security measures to protect personal data, while 8 had conducted privacy impact assessments, and nearly all had set clear data retention policies or allowed users to manage their data.

The review confirmed no violations of the Personal Data (Privacy) Ordinance. In response, the PCPD has laid out several recommendations for organizations using AI, urging them to ensure PDPO compliance, establish clear AI development and use strategies, conduct thorough risk assessments, and engage transparently with stakeholders to refine AI applications and address privacy concerns.



Join the movement & empower the AI policy change-makers

Your contribution makes a real difference:

Provide invaluable AI policy training to aspiring leaders in over 80 countries.

Support the development of our AI & Democratic Values Index Report, a vital resource for understanding AI's global impact.

Stay informed with our comprehensive updates on AI policy developments worldwide.

Help us continue our essential dialogue with policymakers, shaping the responsible evolution of AI.

Every donation helps the Center for AI and Digital Policy build a more informed and equitable AI-driven world.

[Donate with PayPal](#)

Need help with your donation? Contact us at fundraising@caidp.org

CAIDP is a registered non-profit organization, incorporated in Washington, DC

Tax ID: 86-3350258

CAIDP ACTIONS

CAIDP Analysis: Tech Accord Targets AI Misinformation in 2024 Elections

In an era where AI holds the potential to redefine the fabric of democracy, the "[Tech Accord to Combat Deceptive Use of AI in 2024 Elections](#)," announced at the Munich Security Conference, represents an important commitment by some of the world's technology behemoths. This pledge, signed by industry giants including Adobe, Amazon, Google, IBM, Meta, Microsoft, OpenAI, TikTok, and X, aims to confront the burgeoning threat of AI-generated misinformation head-on. However, a closer examination of the accord and the subsequent actions taken by these firms reveals a complex tapestry of commitment and capability, raising critical questions about the collective efficacy of their efforts.



The Accord's Promise and Scope

The accord itself sets a broad yet ambitious agenda, targeting "AI-generated audio, video, and images that deceptively fake or alter the appearance, voice, or actions of political candidates, election officials, and other key stakeholders in a democratic election." It's a direct acknowledgment of AI's potential to undermine electoral processes, emphasizing the development and implementation of technology to mitigate risks related to deceptive AI election content and fostering cross-industry resilience.



Google, through the words of Kent Walker, President of Global Affairs, [underscores](#)

the dual role of AI in both offense and defense within the cyber domain. Google's launch of an AI Cyber Defense Initiative aims to "shore up the cyber defenses of the world's democracies," a commitment that resonates with the accord's call for technological solutions to detect and address online distribution of deceptive AI content. The [initiative](#) proposes a policy and technology agenda to shift the balance from attackers to defenders by enhancing threat detection, malware analysis, vulnerability detection and fixing, and incident response with AI.

Meta's approach, as reflected in a Twitter [post](#) by Nick Clegg, President of Global Affairs, and a [press release](#) about their WhatsApp Helpline in India, suggests a more focused effort on public awareness and direct engagement with misinformation. Clegg's statement, "This work is bigger than any one company & will require a huge effort across industry, government, & civil society," echoes the accord's emphasis on a collective response to the challenge of AI-generated misinformation. Meta's initiative in India, aiming to combat AI-generated misinformation through a dedicated fact-checking helpline, represents a tangible step towards fulfilling the accord's pledge to support efforts to foster public awareness and resilience.

Microsoft's Brad Smith, Vice Chair & President [paints](#) a vivid picture of the stakes involved, stating, "This quickly has become a year that requires all of us who care about democracy to work together to meet the moment." Microsoft's response involves a comprehensive strategy that includes developing tools and services with safety architecture to prevent abuse, alongside fostering transparency and societal resilience.

Adobe's [response](#), as articulated by Dana Rao, Executive Vice President, highlights a commitment to transparency and content authenticity, aiming to "restore trust in digital content." Rao points to the importance of "developing and implementing technological solutions like provenance, watermarking, and classifiers to label and provide metadata about AI-generated content."

CAIDP's Rotenberg Speaks at World Forum on the Future of Democracy, Tech, and Humankind

During the 2024 Cinema for Peace Foundation event in Berlin, Marc Rotenberg, Executive Director of CAIDP, delivered a presentation at the World Forum on the Future of Democracy, Tech, and Humankind. The forum, integral to an event celebrating the power of film to influence global challenges, provided a platform for discussions on democracy,



technology, and humanity's future. Rotenberg emphasized the significance of democratic values in the AI era and highlighted the vital role of the youth in navigating upcoming challenges. For more insights into Marc's reflections, follow the link to his [LinkedIn post](#).

Promise vs. Practice

Yet, the accord and the accompanying corporate pledges evoke critical questions about efficacy, accountability, and the very nature of collaboration in the digital age. The initiatives, as diverse as they are, reveal a patchwork of approaches to a problem that is fundamentally borderless and evasive. Adobe's and Google's focus on technological fixes, Meta's regional initiatives, Microsoft's comprehensive strategy, and the broader, somewhat nebulous commitments of others like OpenAI, illustrate the complexity of operationalizing such an accord.

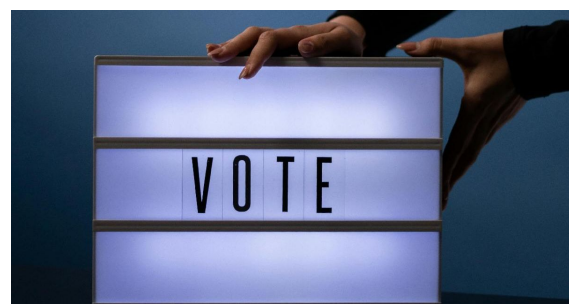


The lack of clear metrics for success and accountability mechanisms within the accord itself becomes a glaring oversight when juxtaposed with the ambitious goals set forth. While the accord laudably aims to "deploy technology countering harmful AI-generated content meant to deceive voters," the pathway to tangible outcomes remains obscured by the very nature of the technology it seeks to regulate. AI's rapid evolution and its capacity for misuse in an election context demand not just a commitment to action but a framework for measurable impact and cross-jurisdictional cooperation.

In an earlier [statement](#), CAIDP urged the Federal Election Commission (FEC) in the U.S. to clarify regulations to expressly include deceptive AI campaign ads. The Center emphasized the "unique risks of generative AI and machine learning systems," distinguishing them from other technologies due to their capacity for producing "indistinguishable falsehoods." This critical viewpoint sheds light on the tech industry's self-regulated efforts, questioning whether they sufficiently address the broader implications of AI on election integrity.

A Call for Continued Vigilance and Collaboration

The CAIDP's recommendations to the FEC highlight a crucial gap between the



tech industry's internal safeguards and the need for external, enforceable standards. By advocating for amendments to FEC regulations, the CAIDP seeks to establish a legal framework that holds political and commercial actors accountable for the misuse of AI in campaign materials. This approach underscores a fundamental principle: transparency is not merely a voluntary commitment but a regulatory requirement to safeguard democracy.

In essence, the Tech Accord marks a significant step towards safeguarding democracy in the digital age. The commitments outlined by its signatories offer a glimpse into the potential for collective action against a pervasive threat. Yet, as the digital landscape continues to evolve, the real test of the accord's legacy will lie not in the commitments themselves but in their execution and the collaborative spirit they engender. As we march towards a future where the lines between digital content and reality blur, the accord's true success will be measured by its ability to adapt, to hold its signatories accountable, and, ultimately, to preserve the integrity of the world's democracies.

AI POLICY EVENTS

- “How can we finally fix the EU-US privacy quarrel?” with Marc Rotenberg (CAIDP), CEPS Ideas Lab, Brussels, March 4-5
- Committee on Artificial Intelligence Plenary meeting with Karine Caunes (CAIDP), Council of Europe, Strasbourg, March 11, 14, 2024
- CAIDP AI and Democratic Values, Washington, DC, April 4 2024
- CAIDP Graduation with keynote speaker Dr. Alondra Nelson, Washington DC, April 8, 2024
- Project Voice with Merve Hickok (CAIDP), Tennessee, April 22
- World AI Summit America with Merve Hickok (CAIDP), Montreal, April 24-25
- To Govern or to be Governed, that is the Question, CPDP.AI, Brussels, May 22-24, 2024



17th international conference
22-24 May ■ Brussels, Belgium
TO GOVERN OR TO BE GOVERNED,
THAT IS THE QUESTION

[Subscribe](#) to CAIDP | Let us know your suggestions [here](#).
| [Privacy Policy](#) | [Contact](#) | [Donate](#) |



CAIDP | 1100 13th ST NW, Suite 800, Washington, DC 20005

[Unsubscribe rotenberg@caidp.org](mailto:Unsubscribe_rotenberg@caidp.org)

[Update Profile](#) | [Constant Contact Data Notice](#)

Sent by admin@caidp.ccsend.com powered by



Try email marketing for free today!