Center for AI and
Digital Policy

October 18, 2023

Chairman Cathy McMorris Rodgers and Ranking Member Frank Pallone
Chairman Chair Gus Bilirakis and Ranking Member Jan Schakowsky
House Committee on Energy and Commerce
Subcommittee on Innovation, Data, and Commerce
2125 Rayburn House Office Building
Washington D.C. 20515

Re: CAIDP Statement for the Record: *"Safeguarding Data and Innovation: Building the Foundation for the Use of Artificial Intelligence", October 18, 2023*

Dear Chairman Rodgers and Ranking Member Pallone, Chairman Bilirakis and Ranking Member Schakowsky, and Committee Members,

We write to you, on behalf of the Center for AI and Digital Policy (CAIDP)[1], regarding today's hearing on *"Safeguarding Data and Innovation: Building the Foundation for the Use of Artificial Intelligence."*[2]

We previously testified on AI policy before the House Oversight Committee on *Advances in AI: Are We Ready For a Tech Revolution?*[3] As CAIDP President Merve Hickok told the Committee: "We do not have the guardrails in place, the laws that we need, the public education, or the expertise in government to manage the consequences of the rapid changes that are now taking place."[4]

[1] Center for AI and Digital Policy, https://www.caidp.org

[2] House Energy & Commerce Committee, Subcommittee on Innovation, Data, and Commerce Hearing, *Safeguarding Data and Innovation: Setting the Foundation for the use of Artificial Intelligence*, Oct. 18, 2023, https://energycommerce.house.gov/events/safeguarding-data-and-innovation-setting-the-foundation-for-the-use-of-artificial-intelligence

[3] Testimony and statement for the record of CAIDP President Merve Hickok, *Advances in AI: Are We Ready For a Tech Revolution?* House Committee on Oversight and Accountability: Subcommittee on Cybersecurity, Information Technology, and Government Innovation, Mar. 8, 2023, https://oversight.house.gov/wp-content/uploads/2023/03/Merve-Hickok_testimony_March-8th-2023.pdf.

[4] *Id.*

Your hearing on *Safeguarding Data and Innovation* could help address some of these challenges. In this statement, we respond to the questions set out in the hearing memo.[5] Our recommendations in brief are:

1. Congress should move forward with comprehensive AI legislation.

2. The FTC should complete its investigation of OpenAI with urgency

3. Congress should promote the use of Privacy Enhancing Technologies (PETs) that minimize or eliminate the collection and use of personal data

<u>About CAIDP</u>

The Center for AI and Digital Policy (CAIDP) is an independent research organization based in Washington, DC. We advise national governments and international organizations regarding artificial intelligence and digital policy. CAIDP currently serves as an advisor on AI policy to the OECD, the Global Partnership on AI, the European Union, the Council of Europe, UNESCO, and other national and international organizations. In April 2023, we released the third edition of our *Artificial Intelligence and Democratic Values* Index, providing a comprehensive review of AI policies and practices in 75 countries.[6]

1. What privacy concerns arise from the widespread adoption of large language models and other forms of artificial intelligence?

In a complaint filed earlier this year with the Federal Trade Commission, CAIDP warned of the specific Risks to Privacy of generative AI products such as ChatGPT.[7] OpenAI itself has acknowledged that "GPT-4 has the potential to be used to attempt to identify private individuals when augmented with outside data."[8]

ChatGPT exposes users to risk of loss of data and privacy violations through "Conversational AI leaks."[9] Several data breaches were reported after the deployment and

[5] Subcommittee on Innovation, Data, and Commerce, *Safeguarding Data and Innovation: Building the Foundation for the Use of Artificial Intelligence,* https://d1dth6e84htgma.cloudfront.net/IDC_Hearing_Memo_on_AI_101823_1_76e2b78b7b.pdf

[6] CAIDP, *Artificial Intelligence and Democratic Values* (2023), https://www.caidp.org/reports/aidv-2022/

[7] CAIDP, *In the Matter of OpenAI* (2023), https://www.caidp.org/cases/openai/.

[8] *Id* at pg. 22

[9] CAIDP, *Supplement to the Original Complaint, In the Matter of Open AI* (2023), pg. 20, para. 85, https://files.constantcontact.com/dfc91b20901/72cccde7-44a7-44e4-bfee-d6801b3891d2.pdf. *See also,*

incorporation of ChatGPT in consumer facing services and applications.[10] Companies such as Samsung, Bank of America, Goldman Sachs, Wells Fargo, Citigroup, and Deutsche Bank have imposed internal restrictions on the use of ChatGPT and other generative AI tools.[11]

In a study commissioned by ELSA – European Lighthouse on Secure and Safe AI - computer security experts concluded that Large Language Models (LLMs) "blur the line between data and instructions" due to security flaws of "indirect prompt injections" and render LLMs such as ChatGPT vulnerable to "data theft, worming, information ecosystem contamination, and other novel security risks." They urged ongoing security evaluations of LLM-integrated applications.[12]

In June 2023, the Norwegian Consumer Council (NCC) published *Ghost in the Machine*, a report on the consumer dangers of generative AI.[13].The NCC report presents several critical findings:

a) "Generative AI models are dependent on large amounts of data that is taken from a multitude of sources, usually without the knowledge or consent of the originator of the data, be it a piece of art, a news article, or a selfie.
b) There are rising concerns about generative AI in chat- bots and their ability to trick consumers into sharing personal data, which may be repurposed to serve targeted advertising or to manipulate consumers into purchasing products or services. This is especially relevant in the case of vulnerable groups such as children or lonely people, who may be more likely to share sensitive information about themselves in conversation with the generative AI.
c) Image generators are usually trained on huge datasets that include images of real people. These images can, for example, be taken from social media and search engines,

Vilius Petkauskas, *Lessons learned from ChatGPT's Samsung Leak,* Cybernews, May 9, 2023, https://cybernews.com/security/chatgpt-samsung-leak-explained-lessons/

[10] Stefanie Schappert, *ChatGPT leaks user credit card details,* Cybernews, Jun. 2, 2023, https://cybernews.com/news/payment-info-leaked-openai-chatgpt-outage/; *See also,* Abraham Jewett, *ChatGPT allegedly suffers outage, data breach,* Top Class Actions, Apr. 4, 2023, https://topclassactions.com/lawsuit-settlements/privacy/data-breach/chatgpt-allegedly-suffers- outage-data-breach/; Sue Poremba, *ChatGPT Confirms Data Breach, Raising Security Concerns,* Security Intelligence, May 2, 2023, https://securityintelligence.com/articles/chatgpt-confirms- data-breach/

[11] Siladitya Ray, *Samsung bans ChatGPT Among Employees After Sensitive Code Leak,* Forbes, May 2, 2023, https://www.forbes.com/sites/siladityaray/2023/05/02/samsung-bans-chatgpt- and-other-chatbots-for-employees-after-sensitive-code-leak/?sh=6f59bcb96078; Gabriela Mello, William Shaw and Hannah Levitt, *Wall Street Banks Are Cracking Down on AI-Powered ChatGPT*, Bloomberg News, Feb. 24, 2023, https://www.bloomberg.com/news/articles/2023-02-24/citigroup-goldman-sachs-join-chatgpt-crackdown-fn-reports?sref=CSMHWBLp#xj4y7vzkg

[12] *Id* at pg. 18, para. 79

[13] *Id* at pg. 28, 29, para. 110

> without a lawful legal basis or knowledge by the people in the pictures. Similarly, text generators are trained data-imasets that could include personal data about individuals, or conversations between individuals."[14]

In April 2023, Snap Inc. integrated ChatGPT into Snapchat – a popular application used by many children. There is no opt-out from the chatbot in the standard, free version of the app. The information page on Snapchat states "My AI "may include biased, incorrect, harmful or misleading content" and suggests that users should independently verify any advice it gives before acting on it."[15] Snap's decision to integrate ChatGPT into applications widely used by children will enable further collection of children's personal data as kids may not be aware of the nature of the information they are disclosing to the chatbot.[16]

Just last month, OpenAI released the third version of DALL-E, a software for AI-generated art.[17] Now, DALLE-3 incorporates ChatGPT.[18] OpenAI has announced "voice and image capabilities in ChatGPT."[19] OpenAI has announced its vocal technology will be "capable of crafting realistic synthetic voices from just a few seconds of real speech."[20]

The use of image-to-text techniques to analyze images of people has staggering implications for personal privacy and personal autonomy, as it would give the user of GPT-4 the ability not only to link an image of a person to detailed personal data, available in the model, but also for OpenAI's product GPT-4 to make recommendations and assessments, in a conversational manner, regarding the person.[21]

[14] Norwegian Consumer Council, *Ghost in the machine – Addressing the consumer harms of generative AI*, Jun. 2023, https://storage02.forbrukerradet.no/media/2023/06/generative-ai- rapport-2023.pdf

[15] Bernard Marr, *Snapchat Debuts ChatGPT – Powered Snap AI: But is it safe for kids?*, Forbes, Apr. 26, 2023, https://www.forbes.com/sites/bernardmarr/2023/04/26/snapchat-debuts-chatgpt- powered-snap-ai--but-is-it-safe-for-kids

[16] CAIDP, *Supplement to the Original Complaint, In the Matter of Open AI* (2023), pg. 25, 26, para. 105, 106, https://files.constantcontact.com/dfc91b20901/72cccde7-44a7-44e4-bfee-d6801b3891d2.pdf

[17] Emilia David, *OpenAI releases third version of DALL-E*, The Verge, Sept. 20, 2023 https://www.theverge.com/2023/9/20/23882009/class-action-lawsuit-openai-privacy-dropped.

[18] Will Knight, *OpenAI's Dall-E 3 is an Art Generator Powered by ChatGPT*, Wired, Sept. 20, 2023, https://www.wired.com/story/dall-e-3-open-ai-chat-gpt/.

[19] OpenAI Blog, *ChatGPT can now see, hear and speak,* Sept. 25, 2023, https://openai.com/blog/chatgpt-can-now-see-hear-and-speak.

[20] TechCrunch, *OpenAI gives ChatGPT a voice for verbal conversations,* Sept. 25, 2023, https://techcrunch.com/2023/09/25/openai-chatgpt-voice/.

[21] CAIDP, *Supplement to the Original Complaint, In the Matter of Open AI* (2023), pg. 25, 26 para. 105, 106, https://files.constantcontact.com/dfc91b20901/72cccde7-44a7-44e4-bfee-d6801b3891d2.pdf

AI systems can replicate and amplify human biases. NIST has also reported on the problems of bias in data sets.[22] Further, vulnerable populations are the most affected by biased AI outputs. There are examples of datasets created from human biases influencing AI to replicate systemic racism in facial recognition.[23]

The Federal Trade Commission (FTC) should move forward and conclude its investigation of OpenAI initiated by CAIDP. Subsequent to the filing of the CAIDP Complaint regarding OpenAI, consumer agencies around the world have launched investigations of ChatGPT.[24] The Federal Trade Commission (FTC) now has a unique opportunity at this time to establish guardrails for AI. Earleir this year, CAIDP filed a detailed complaint with the FTC regarding OpenAI. The FTC has opened the investigation of OpenAI we requested.[25] This is clearly a positive development, but now the FTC must prioritize this investigation. It took two years from the time we filed similar complaints with the FTC concerning Google and Facebook before there was a settlement.[26] We can't wait that long this time. AI products are evolving rapidly and being deployed downstream in consumer facing services. Before the end of this year the FTC must complete the investigation of ChatGPT and enter into a settlement with OpenAI that ensures the companies will abide by the practices for AI companies the Commission has previously issued.

2. How can increasing user control over their data help alleviate concerns about the future of AI?

Congress must move forward with federal AI legislation. We need federal AI legislation which would mandate transparency, fairness, and accountability of AI systems. As CAIDP President Merve Hickok said in her testimony, "AI systems determine people's opportunities in life – from education and credit to employment and housing. The use of opaque and unprovable

[22] NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY, *More to AI Bias than Biased Data, NIST Report Highlights*, Mar. 1, 2022, https://www.nist.gov/news-events/news/2022/03/theres-more-ai-bias-biased-data-nist-report-highlights.

[23] *Id*.

[24] CAIDP, *Supplement to the Original Complaint, In the Matter of Open AI* (2023), pg. 2, para. 6, https://files.constantcontact.com/dfc91b20901/72cccde7-44a7-44e4-bfee-d6801b3891d2.pdf.

[25] Cecilia Kang and Cade Metz, *F.T.C. Opens Investigation Into ChatGPT Maker Over Technology's Potential Harms*, The New York Times, Jul. 13, 2023, https://www.nytimes.com/2023/07/13/technology/chatgpt- investigation-ftc-openai.html; John D. McKinnon and Ryan Tracy, *ChatGPT Comes Under Investigation by Federal Trade Commission,* Wall Street Journal, Jul. 13, 2023, https://www.wsj.com/articles/chatgpt-under-investigation-by- ftc-21e4b3ef.

[26] Federal Trade Commission, *Facebook Settles FTC Charges That It Deceived Consumers By Failing To Keep Privacy Promises*, Nov. 29, 2011, https://www.ftc.gov/news-events/news/press-releases/2011/11/facebook-settles- ftc-charges-it-deceived-consumers-failing-keep-privacy-promises

decision-making systems means that both people subject to the systems and the companies that deploy systems do not actually know the basis of outcomes."[27]

For this reason, developers of foundation models as well as downstream users should be held responsible for ensuring safety by design and implementing specific safeguards on transparency of data practices and disclosure on AI-generated content.

Leading AI researchers, Emily Bender, Timnit Gebru, and Margaret Mitchell have explained, "researchers must adopt frameworks to describe the uses for which their models are suited and benchmark evaluations for a variety of conditions. This involves providing thorough documentation on the data used in model building, including the motivations underlying data selection and collection processes. This documentation should reflect and indicate researchers' goals, values, and motivations in assembling data and creating a given model."[28]

We endorse the Hawley-Blumenthal bipartisan AI Act which offers a comprehensive framework for the governance of AI.

Many machine learning systems rely on the collection of personal data. These systems can violate privacy and be used for mass surveillance. New AI systems pose new challenges for privacy, dignity, autonomy, and equality. Metrics for explainability, interpretability, and transparency should be established to protect fundamental rights, human well-being, and to increase public trust. These metrics alongside Privacy Enhancing Technologies would help protect privacy.[29] **Privacy Enhancing Techniques (PETs) minimize or eliminate the collection and use of personal data.**

There are other techniques, such as 'Privacy by Design' and 'Privacy- Preserving Technologies.' However, sharing data even in anonymized form also presents significant privacy and security risks. Policymakers could support the development and adoption of solutions that

[27] Testimony and statement for the record of CAIDP President Merve Hickok, *Advances in AI: Are We Ready For a Tech Revolution?,* House Committee on Oversight and Accountability: Subcommittee on Cybersecurity, Information Technology, and Government Innovation, Mar. 8, 2023, pg. 4, https://oversight.house.gov/wp-content/uploads/2023/03/Merve-Hickok_testimony_March-8th-2023.pdf.

[28] Emily M. Bender, Timnit Gebru, Angelina McMillan-Major,
Margaret Mitchell, *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big,* FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, March 2021, Pg. 610-623, https://doi.org/10.1145/3442188.3445922

[29] CAIDP, *Comments to* OSTP *on National Artificial Intelligence Research and Development Strategic Plan,* Mar. 4, 2022, Pg. 4, https://www.caidp.org/statements/

allow for secure, responsible use of these data sets, such as privacy-enhancing technologies (PETs).[30]

Where it is necessary to transfer personal data, the most robust methods should be adopted. The adoption of PETs has the potential to address both the risks and opportunities for innovation and robust privacy protections.

Thank you for your consideration of our views. We would be pleased to provide you and your staff with additional information. We ask that this statement be included in the hearing record.

Sincerely,

Merve Hickok
CAIDP President

Marc Rotenberg
CAIDP Executive Director

Christabel Randolph
CAIDP Law Fellow

[30] Latanya Sweeney, Michael von Loewenfeldt, and Melissa Perry, *Saying it's Anonymous Doesn't Make It So: Re-identifications of "anonymized" law school data,* Technology Science, 2018111301, Nov. 12, 2018, https://techscience.org/a/2018111301/; Latanya Sweeney, *Weaving Technology and Policy Together to Maintain Confidentiality*, The Journal of Law, Medicine & Ethics, Vol. 25, Issue. 2-3, pg. 98-110, https://doi.org/10.1111/j.1748-720X.1997.tb01885.x; The Harvard Gazette, *You're not so anonymous,* Oct.18, 2011, https://news.harvard.edu/gazette/story/2011/10/youre-not-so-anonymous/

Center for AI and
Digital Policy

Center for AI and
Digital Policy

ARTIFICIAL INTELLIGENCE AND DEMOCRATIC VALUES INDEX

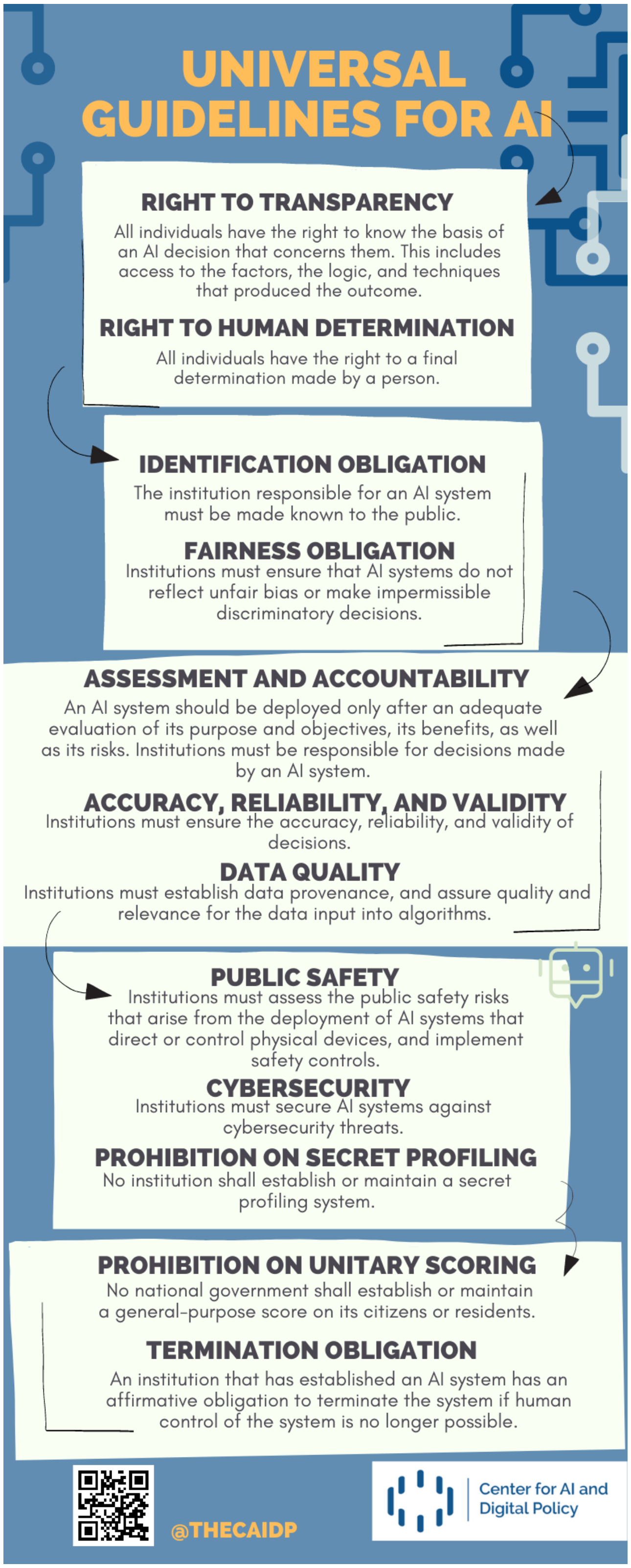
APRIL, 2023

CENTER FOR AI AND DIGITAL POLICY

WASHINGTON, DC

CAIDP.ORG

Center for AI and
Digital Policy



UNIVERSAL GUIDELINES FOR AI

RIGHT TO TRANSPARENCY

All individuals have the right to know the basis of an AI decision that concerns them. This includes access to the factors, the logic, and techniques that produced the outcome.

RIGHT TO HUMAN DETERMINATION

All individuals have the right to a final determination made by a person.

IDENTIFICATION OBLIGATION

The institution responsible for an AI system must be made known to the public.

FAIRNESS OBLIGATION

Institutions must ensure that AI systems do not reflect unfair bias or make impermissible discriminatory decisions.

ASSESSMENT AND ACCOUNTABILITY

An AI system should be deployed only after an adequate evaluation of its purpose and objectives, its benefits, as well as its risks. Institutions must be responsible for decisions made by an AI system.

ACCURACY, RELIABILITY, AND VALIDITY

Institutions must ensure the accuracy, reliability, and validity of decisions.

DATA QUALITY

Institutions must establish data provenance, and assure quality and relevance for the data input into algorithms.

PUBLIC SAFETY

Institutions must assess the public safety risks that arise from the deployment of AI systems that direct or control physical devices, and implement safety controls.

CYBERSECURITY

Institutions must secure AI systems against cybersecurity threats.

PROHIBITION ON SECRET PROFILING

No institution shall establish or maintain a secret profiling system.

PROHIBITION ON UNITARY SCORING

No national government shall establish or maintain a general-purpose score on its citizens or residents.

TERMINATION OBLIGATION

An institution that has established an AI system has an affirmative obligation to terminate the system if human control of the system is no longer possible.

@THECAIDP

Center for AI and
Digital Policy