January 2024

2024 AI POLICY FORECAST

Gregory C. Allen
Georgia Adamson

CSIS | WADHWANI CENTER FOR AI AND ADVANCED TECHNOLOGIES

A Report of the Wadhwani Center for AI and Advanced Technologies

January 2024

2024 AI POLICY FORECAST

Gregory C. Allen
Georgia Adamson

A Report of the Wadhwani Center
for AI and Advanced Technologies

CSIS | WADHWANI CENTER FOR AI
AND ADVANCED TECHNOLOGIES

ABOUT

CSIS

The Center for Strategic and International Studies (CSIS) is a bipartisan, nonprofit policy research organization dedicated to advancing practical ideas to address the world's greatest challenges.

Thomas J. Pritzker was named chairman of the CSIS Board of Trustees in 2015, succeeding former U.S. senator Sam Nunn (D-GA). Founded in 1962, CSIS is led by John J. Hamre, who has served as president and chief executive officer since 2000.

CSIS's purpose is to define the future of national security. We are guided by a distinct set of values—nonpartisanship, independent thought, innovative thinking, cross-disciplinary scholarship, integrity and professionalism, and talent development. CSIS's values work in concert toward the goal of making real-world impact.

CSIS scholars bring their policy expertise, judgment, and robust networks to their research, analysis, and recommendations. We organize conferences, publish, lecture, and make media appearances that aim to increase the knowledge, awareness, and salience of policy issues with relevant stakeholders and the interested public.

CSIS has impact when our research helps to inform the decisionmaking of key policymakers and the thinking of key influencers. We work toward a vision of a safer and more prosperous world.

CSIS does not take specific policy positions; accordingly, all views expressed herein should be understood to be solely those of the author(s).

THE WADHWANI CENTER FOR AI AND ADVANCED TECHNOLOGIES

The Wadhwani Center for AI and Advanced Technologies is an initiative within CSIS that produces research on technology governance, regulation, national security, and geopolitics—with a particular focus on AI. The center investigates central topics including export controls, semiconductor supply chains, and the impacts of emerging technologies on national security and global economic policymaking. Our analyses aim to inform policy solutions that address rapidly evolving technology, foster global collaboration on AI and technology, and encourage technological innovation. The center is supported by the Wadhwani Foundation, a nonprofit dedicated to accelerating economic development.

The Wadhwani Center for AI and Advanced Technologies would like to thank the Wadhwani Foundation for making this report possible.

CONTENTS

LETTER FROM THE DIRECTOR

2023 marked the founding year of the Wadhwani Center for AI and Advanced Technologies. We established this organization at CSIS to provide impactful research and analysis that can keep pace with the rapidly changing technology and policy landscape. While this is always a struggle, I am immensely proud of the strides we have made.

In just the eight months that the Wadhwani Center has been active, we proudly published 10 comprehensive reports, offering insights and recommendations that have resonated with policymakers and industry leaders alike. Our experts testified before Congress on three occasions and had the privilege of briefing our research findings to senior international and U.S. government policymakers, including cabinet-level officials, on dozens of occasions.

Looking ahead, we remain committed to advancing the discourse on AI and technology. Our goals for 2024 are ambitious, focusing on growing our team of talented staff, deepening our research, expanding our outreach, convening timely events, and continuing to inform policy at both national and international levels.

I would like to extend a special thank you to Dr. Romesh Wadhwani and the Wadhwani Foundation for their unwavering and generous support. Their commitment to independent and impactful policy scholarship gave us this opportunity. I am deeply grateful to all our donors, partners, and staff who have worked so hard to get us off to such a strong start.

Thank you for being part of our story. I am confident that this year will bring even more exciting opportunities.

Sincerely,

Gregory C. Allen

Director, Wadhwani Center for AI and
Advanced Technologies

2023 YEAR IN REVIEW

A Timeline of Major Developments in AI in 2023

WIN MCNAMEE/GETTY IMAGES

JAN

10

China's Cyberspace Administration implements a new law to manage deepfakes, including by enforcing watermarks on AI-generated content.[1]

23

Microsoft pledges a rumored $10 billion multiyear investment in OpenAI, claiming the future impact of AI technology will be equal to the PC or the internet.[2]

25

The U.S. Department of Defense updates DOD Directive 3000.09, "Autonomy In Weapons Systems," changing the original 2012 directive to reflect new technology capabilities in autonomous systems and AI.[3]

26

The U.S. National Institute of Standards and Technology (NIST) releases the NIST AI Risk Management Framework, a set of guidelines for AI development, use, and evaluation aimed to enhance transparency and security of AI in businesses and organizations.[4]

FEB

27

The United States and the European Union announce an agreement to accelerate joint AI research for solving global challenges in climate forecasting, agriculture, healthcare, critical infrastructure, and more.[5]

Bloomberg reports that the Netherlands and Japan will join U.S. efforts to restrict exports of semiconductor manufacturing equipment to China.[6]

31

The White House launches the U.S.-India Initiative on Critical and Emerging Technologies (iCET), a partnership with India to advance technology and defense research and innovation and to ensure semiconductor supply chain resiliency.[7]

ALEX WONG/GETTY IMAGES

JAAP ARRIENS/NURPHOTO VIA GETTY IMAGES

MAR

2

OpenAI's chatbot, ChatGPT, becomes the fastest-growing application in history, with 100 million users in January.[8]

16

The Department of Defense first announces its "Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy" at the Responsible AI in the Military Domain Summit in The Hague, Netherlands, the first summit of its kind.[9]

24

Meta's large language model (LLM) "LLaMA" is announced for limited release.[10]

8

The Netherlands announces plans to join the United States to restrict semiconductor technology to China.[11]

14

OpenAI reveals its latest large multimodal model ChatGPT-4, which greatly outperforms GPT-3 and other available models in several areas.[12]

21

Nvidia reports it has modified one of its top semiconductor chips, the H100, for export to China as H800 following updated U.S. export controls last year.[13]

22

An open letter calling for pause to all frontier AI developments for six months due to potential catastrophic risk to society is published. Signatories include prominent CEOs and academics.[14]

31

Japan announces it will join the United States and the Netherlands in restricting exports of semiconductor manufacturing equipment overseas.[15]

APR

11

China's Cyberspace Administration reveals draft steps to manage generative AI content, including bringing content in line with China's core values.[16]

30

The G7 concludes Digital and Tech Ministers' Meeting (April 29–30) in Takasaki, Japan, declaring member states' commitment to an internationally cooperative, adaptable, and risk-based approach to AI governance.[17]

MARIO TAMA/GETTY IMAGES

MAY

1

Hollywood writers begin months-long strikes over issues including AI's role in the creative industry.[18]

4

CEOs of top AI developers OpenAI, Anthropic, Microsoft, and Alphabet meet with President Biden to discuss responsible AI innovation, including companies' responsibility to make products safe.[19]

16

OpenAI CEO Sam Altman and IBM vice president Christina Montgomery testify before Congress on the risks of rapid AI development following the quick rise of ChatGPT.[20]

19

G7 leaders gathered in Hiroshima discuss inclusive governance for AI at the 2023 summit.[21]

CSIS

JUN

30

The nonprofit Center for AI Safety publishes a one-sentence statement arguing that mitigating extinction from AI should be a "global priority," which is signed by top AI CEOs and developers, academics, and other civil society figures.[22]

Nvidia briefly joins tech giants in trillion-dollar market valuation, with shares up over 200 percent since late 2022.[23]

STEPHEN ROUSSEAU/GETTY IMAGES

21

Senate Majority Leader Chuck Schumer announces SAFE Innovation Framework for AI Policy at CSIS.[24]

22

The Department of Commerce announces new public working group to implement and build upon NIST's AI Risk Management Framework created in January.[25]

28

OpenAI sued by authors for copyright infringement after training ChatGPT on their works without proper licensing, raising wider copyright concerns around AI systems.[26]

ED JONES/GETTY IMAGES

CHIP SOMODEVILLA/GETTY IMAGES

CHIP SOMODEVILLA/GETTY IMAGES

CHRIS J. RATCLIFFE/GETTY IMAGES

JUL

18
UN Security Council convenes to discuss AI risks for first time.[27]

Meta launches open-source LLM "LLaMA 2."[28]

21
Leading AI firms including OpenAI, Meta, and Google make voluntary commitments to the White House for ensuring safe AI, including testing products before release and watermarking AI-generated content.[29]

24
Japan's export controls on 23 types of semiconductor manufacturing equipment go into effect.[30]

26
OpenAI, Anthropic, Google, and Microsoft announce new industry body Frontier Model Forum, founded to promote the responsible development of AI and to share knowledge with policymakers.[31]

AUG

8
Nvidia announces new cutting-edge semiconductor chip GH200, speeding up processing times for generative AI systems.[32]

9
President Biden signs executive order banning U.S. investment in sensitive technologies such as AI in China for national security and competition reasons.[33]

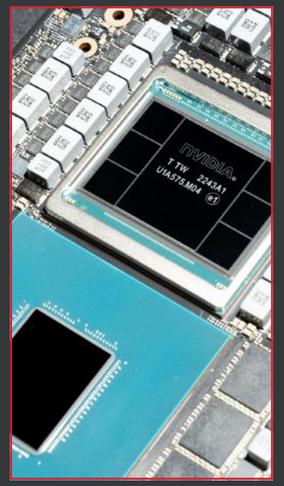
10
The Department of Defense reveals new AI taskforce "Lima" to oversee integration of generative AI capabilities into the department.[34]

28
The Department of Defense unveils new Replicator initiative to accelerate procurement and fielding of all-domain autonomous and attritable military systems to compete with China.[35]

SEP

29
Huawei releases new Mate60 Pro smartphone with 7-nanometer semiconductor chip, highlighting China's technical advancements despite U.S. export and investment restrictions.[36]

31
Chinese tech company Baidu releases AI chatbot Ernie Bot to the public in China.[37]



MARLENA SLOSS/BLOOMBERG VIA GETTY IMAGES

OCT

1
New Dutch restrictions on exporting semiconductor manufacturing equipment to China go into effect.[38]

13
First AI Insight Forum session is held on Capitol Hill. Led by Senate Majority Leader Chuck Schumer, the meeting convenes senators, prominent tech CEOs, and civil society figures to discuss U.S. government oversight of AI.[39]

28
The National Security Agency announces a new body, the AI Security Center, to oversee AI adoption into U.S. national security systems.[40]

CFOTO/FUTURE PUBLISHING VIA GETTY IMAGES

NOV

9
China aims to boost its total computing power by 50 percent by 2025, Chinese ministry reports.[41]

17
The White House announces new and updated measures to restrict AI and semiconductor technology exports to China and other countries, closing loopholes in 2022 control policies.[42]

26
China accepts UK invitation to take part in the UK AI Safety Summit in November amid controversy.[43]

UK government reveals plan to form the world's first institute for AI safety.[44]

30
President Biden signs Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence.[45]

The G7 releases a statement on the Hiroshima AI Process on AI risks and the benefits of fostering an open environment for global collaboration on AI.[46]

1
The United Kingdom's AI Safety Summit opens at Bletchley Park, convening prominent political and technology leaders to discuss international cooperation in AI governance for two days (November 1–2). Twenty-eight countries and the European Union sign the "Bletchley Declaration," announcing international commitment to AI governance and next steps.[47]

Secretary of Commerce Gina Raimondo announces launch of a U.S. AI safety institute at the UK AI summit.[48]

2
The Department of Defense releases its AI strategy, directing the accelerated adoption of advanced AI within the DOD.[49]

15
Microsoft reveals custom-designed semiconductor chip with aim to cut high costs of AI products.[50]

17
Chief Executive Officer Sam Altman is temporarily ousted from OpenAI by the company's board, only to be reinstated five days later on November 21.[51]

DEC

22
Media source Reuters gains knowledge of new OpenAI project "Q*," rumored to be a breakthrough in developing artificial general intelligence.[52]

24
Russian president Vladimir Putin says Russian AI strategy will be released soon to counter the Western and Chinese monopoly on AI at a conference in Moscow.[53]

24
Meta announces restrictions on AI-generated content for advertising ahead of the 2024 election, including mandatory disclosures of AI-generated advertising to the public.[54]

Researchers discover a weakness in ChatGPT, allowing it to reveal sensitive information by using a single prompt.[55]

30
The U.S. Patent and Trademark Office announces a new Semiconductor Technology Pilot Program designed to encourage innovation in semiconductor manufacturing and support the CHIPS and Science Act.[56]

6
Google launches AI model Gemini, the first model to outperform human experts on Massive Multitask Language Understanding (MMLU).[57]

8
In a global first, the European Union establishes a landmark comprehensive AI regulation in passing the EU AI Act.[58]

13
The *Financial Times* reports that generative AI is widely used by multiple political parties in Bangladesh's 2024 elections as deepfakes and AI-generated misinformation circulate social media and news outlets.[59]

30
The *New York Times* sues Microsoft and OpenAI for copyright infringement, claiming AI chatbots illegally used millions of articles for training.[60]

Nvidia launches advanced gaming chip GeForce RTX 4090 D for Chinese consumers, adapted to comply with updated U.S. export controls.[61]

CSIS

> “Now, friends, we come together at a moment of revolution, not one of weapons or of political power, but a revolution in science and understanding that will change humanity. It’s been said that what the locomotive and electricity did for human muscle a century and a half ago, artificial intelligence is doing for human knowledge today as we speak. But the effect of AI is far more profound and will certainly occur over a much shorter period of time.[62]

U.S. senate majority leader Chuck Schumer
June 21, 2023

2023 TOP TAKEAWAYS

The Wadhwani Center's Key Takeaways from Developments in AI Last Year

Existential risk became a mainstream concern for AI governance.

Though the risk of AI leading to catastrophe or human extinction had been a focus for Elon Musk and many AI researchers in prior years, 2023 saw the issue become a genuine priority among global leaders and government policymakers. The shift was led by calls from high-profile figures in the private sector such as OpenAI CEO Sam Altman, Google DeepMind CEO Demis Hassabis, Tesla and xAI CEO Elon Musk, and AI “godfather” Geoffrey Hinton. In May, a coalition of over 350 leading AI experts and executives signed a one-sentence statement that “mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war.”[63] This statement immediately led to a rhetorical shift among global policymakers.

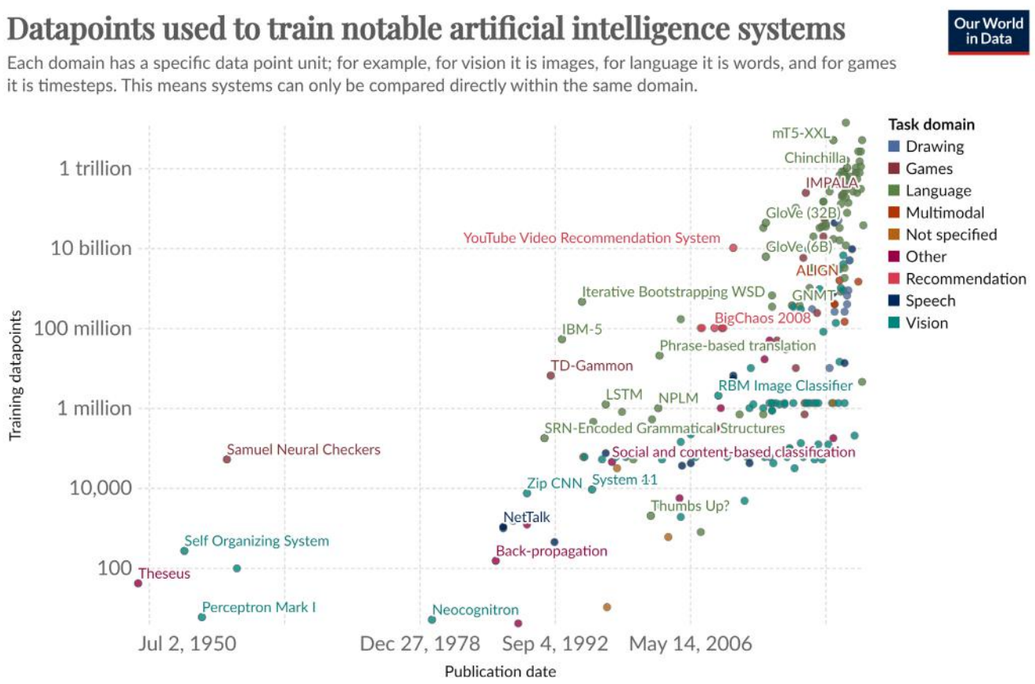
Concern about AI’s malign impact on civilization trickled down to the wider U.S. public. A survey of more than 20,000 Americans by YouGov in April 2023 reported that 46 percent were concerned about AI’s potential to cause human extinction and 69 percent supported a proposed six-month pause on AI development.[64]

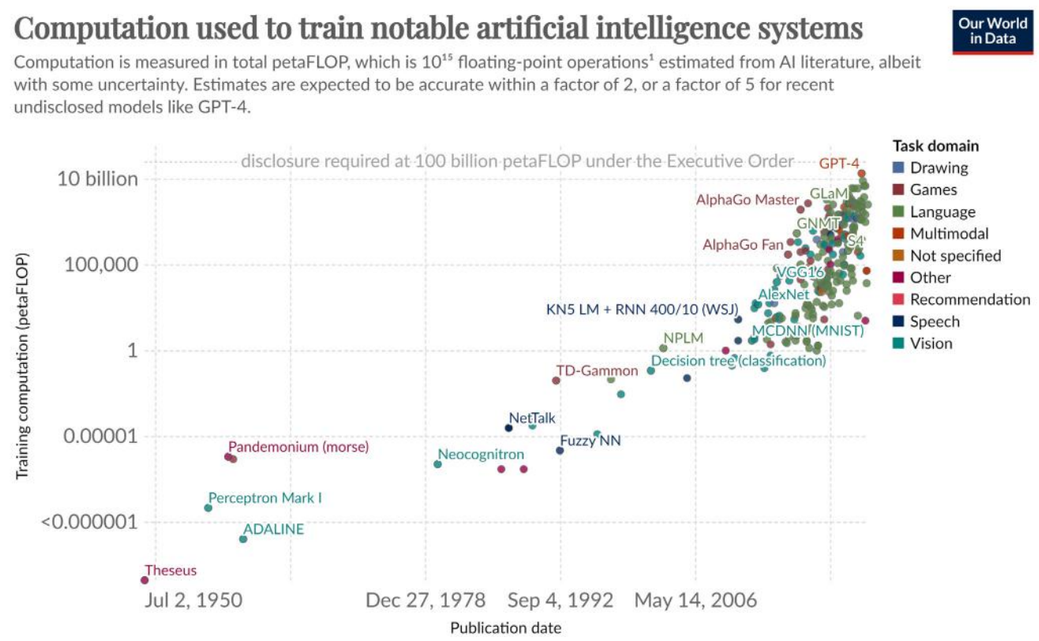
Similar anxieties about the potential catastrophic risk of AI echoed around Washington last year. At a congressional hearing in May, Sam Altman warned that AI could “cause significant harm to the world” and urgently called for greater regulation of the technology.[65] IBM vice president and chief privacy and trust officer Christina Montgomery concurred, “with AI the stakes are simply too high,” and “what we need at this pivotal moment is clear, reasonable policy and sound guardrails.”[66]

At the UK AI Safety Summit, 29 countries attending agreed to the Bletchley Declaration, which stated that “there is potential for serious, even catastrophic, harm, either deliberate or unintentional, stemming from the most significant capabilities of these AI models.”[67]

The United States, the Netherlands, and Japan coordinated export controls to target China’s AI and semiconductor technology development . . .

In late January 2023, reports emerged that the United States had reached an agreement with the Netherlands and Japan for the two countries to impose new export controls restricting China’s access to chipmaking tools.[68] Earlier, on October 7, 2022, the United States had imposed strict controls on exports of advanced semiconductor manufacturing equipment (SME) technology to China; however, as Japanese and Dutch companies produced important SME technology, such unilateral action had a limited effect. Details on what was included in these new restrictions were scarce until March 2023, when the two countries for-





1. **Floating-point operation**: A floating-point operation (FLOP) is a type of computer operation. One FLOP is equivalent to one addition, subtraction, multiplication, or division of two decimal numbers.

mally announced they would be moving forward with export controls on a wide range of semiconductor equipment and technology. Neither country explicitly mentioned China as the target.[69]

Japan and the Netherlands capture 99 percent of the world's market share of lithography steppers and scanners—crucial for state-of-the-art AI chips.[70] Therefore, this was a major step forward in the U.S. mission to bar China from gaining the lead in the chip race. Still, other countries like Germany and South Korea are also significant producers in the semiconductor value chain, and the United States will need to persuade them both to get on board with controls if they want to continue to slow China's technological advancements.

...However, their success in slowing China's technological progress remains mixed.

Despite international efforts to prevent China from making significant technological advancements, the announcement of Huawei's new Mate60 smartphone raised concerns throughout the national security community about the efficacy of the export controls.

On October 17, 2023, the United States' Bureau of Industry and Security announced updates to the October 7 controls.[71] These updates included additional parameters for chips' performance density, the restriction of dozens more items of semiconductor equipment, the expansion of licensing requirements to an additional 22 countries that the United States has an arms embargo with, and the addition of 13 companies to the entity list.[72]

AI chatbots reached billions of users worldwide and continued growing rapidly in scale.

OpenAI's ChatGPT reached over one million users in its first week of launch in November 2022.[73] By November 2023, that number had skyrocketed to more than 100 million monthly active users.[74] OpenAI's success speaks to a wider public entrancement with AI chatbots around the world. Leading AI development companies launched several new large language models (LLMs) last year, including OpenAI's updated ChatGPT-4, Meta's LlaMa 2, and Google's Gemini. The global demand for this technology has prompted many companies to develop chatbots trained on other languages, such as the Arabic model Jais and the Chinese model Ernie Bot, though the United States still leads in the worldwide development of LLMs.[75]

Language models are getting bigger, both in terms of the data they are trained on and their parameters (GPT-4, for example, is rumored to have up to one trillion parameters, compared to GPT-3's 175 billion).[76] In fact, models have grown so large that there are legitimate concerns that companies are reaching the limits of the existing available text training data.[77] However, their growing size comes with growing costs.[78] While training costs are rarely disclosed by companies developing LLMs, OpenAI stated that developing and training GPT-4 cost more than $100 million, and Anthropic CEO Dario Amodei suggested that future training costs could exceed $1 billion.[79] Growing CO2 emissions and water usage from training and operating chatbots are also attracting increased attention in terms of their effect on the environment.[80] What these upward trends mean for AI chatbots' profitability and scalability remains to be seen this year.

Major economies around the world took substantial steps to regulate AI...

The United States

In response to the transformative potential of AI, the U.S. government began to regulate it through administrative law, acknowledging the imperative to navigate the complexities of AI risks and begin establishing domestic standards. On January 26, 2023, the U.S. Department of Commerce's National Institute of Standards and Technology (NIST) unveiled the Artificial Intelligence Risk Management Framework (AI RMF 1.0).[81] Developed in close collaboration with both private and public sectors, the AI RMF serves as a comprehensive tool for organizations engaging with AI technologies and is designed to adapt to the evolving AI landscape. Though the RMF is not intended to be applied as part of formal regulation, many have held the framework up as substantial progress in maturing AI governance.

> We face a genuine inflection point in history, one of those moments where the decisions we make in the very near term are going to set the course for the next decades. And with the position we lead the world, the toughest challenges are the greatest opportunities. Look, there's no greater change that I can think of in my life than AI presents as a potential: exploring the universe, fighting climate change, ending cancer as we know it, and so much more.[82]
>
> U.S. president Joe Biden
> October 30, 2023

Announced in June 2023 at CSIS, Senate Majority Leader Chuck Schumer's SAFE Innovation Framework marked a strategic effort to confront the profound changes brought about by AI through "comprehensive legislation."[83] Since September 2023, Capitol Hill has seen over 150 AI experts gather as part of Senator Schumer's AI Insight Forums.[84] These forums have covered an array of crucial topics, from AI innovation and workforce considerations to national security and guarding against doomsday scenarios.[85] Notably, the ninth forum, held on December 6, 2023, featured a testimony from the Wadhwani Center's director, Gregory Allen.[86] This forum focused on maximizing AI development to bolster the United States' military capabilities, aligning with Senator Schumer's vision for an "all-hands-on-deck" effort.[87]

Ahead of the UK AI Safety Summit, the Biden administration announced its Executive Order on Safe, Secure and Trustworthy Artificial Intelligence in late October 2023.[88] The order broadly focused on the development of standards and testing mechanisms for AI safety, infrastructure, and social consequences (such as discrimination and effects on labor). Since the announcement, executive agencies like the Department of Defense (DOD) and State Department have released their policies on AI and detailed how the executive order's directives will be applied in their respective agencies.[89]

The European Union

On December 8, the European Union passed its Artificial Intelligence Act, the world's most substantial set of regulations on AI so far.[90] After two and a half years in the making, the act was finally agreed upon following a lengthy 37-hour negotiation between EU states and the European Parliament.[91] EU commissioner Thierry Breton confirmed the event on X, stating that "the EU becomes the very first continent to set clear rules for the use of AI" in passing the AI Act and calling it a "launch pad for EU start-ups and researchers to lead the global AI race."[92] Not all EU leaders agree, however; French president Emmanuel Macron condemned the act on December 11, saying "we can decide to regulate much faster and much stronger than our major competitors. But we will regulate things that we will no longer produce or invent. This is never a good idea."[93]

The AI Act regulates all AI sold, used, or deployed within the European Union apart from AI used for military purposes, research, and open-source models, though most provisions only apply to the "high-risk" AI systems. It advances a risk-based approach to managing AI systems by sorting levels of risk into four categories: unacceptable, high, limited, and minimal to none.[94] Unacceptable risks banned under the act include the use of AI for manipulating human behavior, social scoring, and creating biometric databases based on sensitive social categories such as race or religion. The consequences for failing to comply with these rules are steep: under the new rules, companies could be fined €35 million or 7 percent of global revenue.[95]

Full implementation of the new AI law is not expected to begin until 2025 at the earliest, allowing time for the European Commission to establish a regulatory oversight "AI office" in Brussels and companies to adapt to the new rules.[96]

China

In August 2023, China's generative AI measures went into effect.[97] At the time, these measures were some of the most comprehensive regulations on AI and focused on a regulatory framework for generative AI services to the Chinese public. Earlier in the year, China released a draft that placed significant responsibility onto service providers on topics like ensuring the legality of the data used for training and optimization.[98] Service providers would be met with a fine upwards of 100,000 yuan if they failed to meet these standards. However, the finalized version only requires the service providers to create measures that prioritize desired values within data training and optimization. Shifting from the strict first draft to a more diluted final draft is likely a reflection of industry input as well as sensitivity to the current economic challenges that China is facing.

The measures apply to any AI generative technology that provides services that "generate any text, image, audio, video, or other such content to the public."[99] Services that are not offered to the public are explicitly excluded from the legislation. Additionally, prior to providing services to the public, service providers must apply for a security assessment. In implementation, these have not been difficult to get approved.[100] Chinese military, intelligence, and police services remain broadly exempt from all Chinese AI regulations.

In October, the Chinese government announced its Global AI Governance Initiative at the third Belt and Road forum.[101] Xi Jinping unveiled the initiative personally, underscoring China's AI governance ambitions, which include enhancing information exchange and technological cooperation with other countries; developing open, fair, and efficient governing mechanisms; and establishing an international institution within the UN framework to govern AI. The initiative calls for representation and equal rights when developing AI, regardless of a country's "size, strength, or social system."[102] The announcement came just weeks after the creation of a BRICS AI study group which aimed to foster closer AI governance ties among the participating nations.[103]

. . . and to strengthen international cooperation on AI governance.

Global efforts to govern AI increased dramatically in 2023. The United Kingdom hosted the first global AI Safety Summit on November 1–2, convening political and technology leaders for discussions focusing on foundation model safety. Attendees included summit host Prime Minister Rishi Sunak, European Commission president Ursula von der Leyen, Vice President Kamala Harris, prominent tech CEOs, and, to the surprise of many, China's vice minister of science and technology, Wu Zhaohui.[105] The most significant achievement of the summit was the Bletchley Declaration, which Sunak called a "landmark achievement that sees the world's greatest AI powers agree on the urgency behind understanding the risks of AI."[106] The declaration broadly aims to promote transparency and accountability within companies developing frontier AI models and to develop AI safety evaluation metrics, tools, and research capabilities.

> We call for global collaboration to share knowledge and make AI technologies available to the public under open-source terms.[104]
>
> Chinese vice minister of science and technology Wu Zhaohui
> November 1, 2023

AI also featured prominently in G7 meetings in 2023, including the Digital and Technology Ministers' Meeting in April and the G7 summit in May. The group released the G7 leaders' statement, the Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI Systems, and the Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems.[107] The code of conduct was the most consequential of the documents, laying out a set of voluntary guidelines for developing advanced AI systems, such as foundation models and generative AI. It is worth noting that these documents set out voluntary guidelines as opposed to binding international regulations. The G7 countries will also publish the Hiroshima AI Process Comprehensive Policy Framework.

Several other international assemblies gathered in 2023 to discuss AI for the first time. In February, the Responsible AI in the Military Domain Summit (REAIM) in The Hague, Netherlands, convened ministers to discuss the responsible use of AI for military applications.[108] In July, the United Nations Security Council met to discuss AI risk, at which Secretary-General António Guterres called for the creation of a UN body "to support collective efforts to govern this extraordinary technology."[109] The urgent need for international cooperation on balancing potential risks and benefits from AI was a common theme across all these governance efforts.

What remains to be seen is how international pledges will translate into real-world impact in 2024. Last year brought many voluntary commitments and high-level declarations. This is a start. But perhaps the greater challenge will be seeing how the grittier details of legislation, funding, and international standards unfold within a window of interest that may not last forever.

> **The development of AI is as fundamental as the creation of the microprocessor, the personal computer, the Internet, and the mobile phone. It will change the way people work, learn, travel, get health care, and communicate with each other. Entire industries will reorient around it. Businesses will distinguish themselves by how well they use it.[110]**
>
> Philanthropist and Microsoft c o-founder Bill Gates
> March 21, 2023

The private sector emphasized its own role in responsible AI governance.

As the U.S. government made significant steps to begin regulating AI in 2023, the private companies behind AI development have become more, not less, important in pursuing this goal. The government was proactive last year in collaborating with industry leaders to responsibly manage AI. This action has come, in part, from a recognition by Congress that it is playing catch-up to a technology and industry that far predates the government's regulatory efforts in 2023. The majority of AI research and development exists in the private sector, as it takes extremely large datasets, technical expertise, and financial investment to develop the kinds of frontier AI models Congress is seeking to regulate. It would make sense, therefore, that Congress should seek AI companies' input on AI legislation as it does with other industry leaders in other sectors. Senate Majority Leader Chuck Schumer opened the first AI Insight Forum on September 13, which gathered senators, CEOs, and civil society leaders to discuss AI regulation, noting that "Congress cannot do it alone."[111] He added, "We need help of course from developers and experts who build AI systems."[112] The forums accompanied other congressional hearings that heard from AI CEOs like OpenAI's Sam Altman and IBM's Christina Montgomery last year.[113]

In addition to congressional hearings, AI companies were given a greater responsibility to regulate their technologies in 2023 by Biden's Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence.[114] The order requires AI developers conduct red teaming on their products and report their findings to the government before they are released, placing the onus on companies to do the heavy lifting when it comes to due diligence.

However, the private sector's growing influence in AI governance is accompanied by growing concerns that there are fundamental conflicts of interest at play. AI companies' dual role of helping to regulate their innovations as they continue to develop them raises serious questions about whether they can truly place safety over profit. As Marietje Shaake, international policy director at Stanford's Cyber Policy Center, wrote in the *Financial Times*, "Imagine the chief executive of JPMorgan explaining to Congress that because financial products are too complex for lawmakers to understand, banks should decide for themselves how to prevent money laundering, enable fraud detection and set liquidity to loan ratios. He would be laughed out of the room."[115] Just as lawmakers are not experts in other industries for which they craft legislation, she and others argue, they must be careful to not be captured by AI CEOs who use technical complexity to advance their own regulatory interests.

The number of AI-related lawsuits rose sharply.

New AI tools began to test existing legal frameworks in 2023, cutting across domains such as data privacy, copyright, and patent law. So far, the leading cause for lawsuits related to AI is related to data privacy, perhaps unsurprising considering the billions of data points that generative AI models are trained on from across the internet. AI companies have been resistant to revealing their training data thus far, despite long-running calls for greater transparency and concerns about user privacy from academics and civil society groups.[116] A reported hack of ChatGPT in early December highlighted some basis for these concerns: when asked to repeat certain words like "poem" ad infinitum, ChatGPT eventually began to spit out sensitive training data, including phone numbers, names, and addresses.[117] OpenAI has since closed this loophole by restricting ChatGPT from repeating words forever.[118] It also announced in August that website owners can now block the company's data-scraping web crawler, GPTBot, from accessing their pages and data for training purposes.[119]

OpenAI has been one of many AI companies to face several lawsuits in 2023 due to alleged privacy and intellectual property violations. One class action lawsuit against OpenAI and Microsoft made headlines in June 2023 when it claimed that ChatGPT stole millions of sensitive data from hundreds of millions of internet users during training, including from social media accounts, medical records, and personal accounts.[120] Though the complaint was dismissed in court, it was not the only one of its kind; a second lawsuit was filed against the same two companies in September and another, nearly identical, lawsuit was filed against Google parent company Alphabet in July. [121] As of the time of writing, both cases are ongoing, though all three companies have moved to dismiss them in court.

The implications of AI for copyright law have also come under greater scrutiny in the last year. In 2023, Getty Images filed lawsuits in the United States and the United Kingdom against generative AI company Stability AI for training its model, Stable Diffusion, on copyrighted data and metadata.[122] OpenAI, Alphabet, and Microsoft faced multiple class action cases from authors whose work, they claim, has been similarly used for training purposes without proper licensing or compensation.[123] Text-to-image AI generators like Midjourney, DreamStudio, and DreamUp were the subject of a similar lawsuit filed in January by visual artists whose work has been used and replicated without permission.[124]

Several companies, such as Meta, Microsoft, and Google, argued they should not have to pay to train AI models on copyrighted work, citing arguments such as AI training "is like the act of reading a book" and curbing access to copyrighted access would chill AI development.[125] These kinds of arguments ask fundamental questions about AI regulation, like whether the unique characteristics of the technology should exempt it from certain laws.

Finally, the U.S. Federal Court Circuit upheld a precedent in patent law when it ruled in August that AI systems are not eligible to own patents for their "inventions" as they are not human beings.[126] The verdict confirmed earlier rulings made by the U.S. Patent Office and the U.S. Copyright Office following years of legal disputes by U.S. computer scientist Stephen Thaler, who first tried to copyright an image produced by his AI system in 2019. Federal Circuit Judge Leonard Stark announced the decision in August saying "there is no ambiguity: the Patent Act requires that inventors must be natural persons; that is, human beings."[127] The ruling reflects similar decisions made in the United Kingdom and the European Union.[128]

The Department of Defense took steps to safely adopt and deploy AI in its weapons systems . . .

The United States was the first country to codify a policy on autonomy in weapon systems when it first adopted DOD Directive 3000.09 in 2012.[129] The policy did not ban the development or use of autonomous weapons—indeed many types of autonomous weapons, such as some missile defense and cyber weapons, had already been in use for decades. What 3000.09 did was place new policy and process requirements for the development and use of autonomous weapons in offensive and kinetic constructs. However, the policy was widely misunderstood as requiring "a human in the loop" and thus banning fully autonomous systems, which it did not do. In January 2023, the DOD published an updated 3000.09 that sought to address the confusion and to account for the rise in machine learning AI systems.[130] Among other things, it formalized that adherence to the Department of Defense (DOD)'s AI Ethical Principles was a requirement at all stages of development and fielding. Reaffirmed by Deputy Secretary of Defense Kathleen Hicks, the directive mandates rigorous testing, reviews, and senior-level scrutiny for autonomous systems, aligning with the DOD's Ethical Principles and the Responsible AI (RAI) Strategy.[131]

On the international front, the U.S. Bureau of Arms Control, Deterrence, and Stability unveiled a groundbreaking "Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy."[132] This declaration is fully consistent with the ideas underpinning DOD 3000.09 and seeks to make them international, particularly among U.S. allies. Introduced at the Responsible AI in the Military Domain Summit, this initiative has since garnered signatures of support from 49 countries, promoting non-legally-binding guidelines for secure AI deployment in defense contexts.[133]

. . . and to massively accelerate the DOD's adoption of AI-enabled autonomous systems.

On August 28, 2023, Deputy Secretary of Defense Kathleen Hicks announced the Replicator initiative. Replicator aims to "field attritable autonomous systems