The U.S. Plans to 'Lead the Way' on Global Al Policy

By Alan Charles Raul and Alexandra Mushka

Tuesday, February 13, 2024, 8:00AM

*This article was originally published in Lawfare. Link to original article here.

The U.S. has signaled its intent to assert international influence over AI regulation, marking a departure from previous disengagement with data privacy standards.



Kamala Harris attends the Bletchley AI Safety Summit (Number 10, https://www.flickr.com/photos/number10gov/53303905662/in/photostream/, CC BY-NC-ND 2.0 DEED, https://creativecommons.org/licenses/by-nc-nd/2.0/)

Policymakers around the world took significant steps toward regulating artificial intelligence (AI) in 2023. Spurred by the launch of revolutionary large language models such as OpenAI's GPT series of models, debates surrounding the benefits and risks of AI have been brought into the foreground of political thought. Indeed, over the past year, legislative forums, editorial pages, and social media platforms were dominated by AI discourse. And two global races have kicked into high gear: Who will develop and deploy the most cutting-edge, possibly risky AI models, and who will govern them?

In the wake of this competition, it is worth examining whether the United States will yield policy primacy on AI to Europe, or others, as it largely has done in the field of data privacy—or whether it will instead assert leadership on digital governance commensurate with its lead in the digital technology itself. The plethora of federal initiatives adopted in response to the deployment of capable AI systems with significant computational power supports the latter thesis: The United States intends to run ahead of the field on AI governance, analogous to U.S. leadership on cybersecurity rules and governance—and unlike the policy void on privacy that the federal government has allowed the EU to fill. Various policy developments discussed below support this conclusion, chief among them the fact that the aggressive timeline of government action imposed by President Biden's October 2023 AI

executive order means the requirements, imperatives, and guidelines that order sets into motion will almost certainly be in force before the EU's provisional AI Act is adopted and implemented. Indeed, the White House recently announced that every 90-day deadline set forth by the order has been met. Notably, "developers of the most powerful AI systems" are already required "to report vital information" to the Department of Commerce, including the results of safety testing. Nine agencies have submitted risk assessments to the Department of Homeland Security regarding the use of AI in critical infrastructure. The intense level of agency engagement called for by President Biden led the order to be viewed by some in Congress and industry as too powerful, triggering a "campaign to take [it] down," or "defang the industry-facing sections."

The United States' commitment to AI governance is significant given that over the past two decades, global leadership in data-driven technology innovation has become increasingly uncoupled from efforts to regulate that technology. The world's largest data-driven innovators, such as Microsoft, Google, and Meta, are based in the United States. But the world's leading data regulators are based in the European Union—for instance, the apex privacy framework, the General Data Protection Regulation (GDPR), was promulgated by the EU. By comparison, the United States has applied a relatively light hand to the regulation of social media and search. Unlike the European Parliament, Congress has passed no comprehensive law that directly touches these data-driven platforms. Indeed, in 1996, when Congress last stepped in, it was to enact a key liability shield (codified as Section 230 of the Communications Decency Act) that allowed social media companies to grow.

In the text of Section 230, Congress consciously chose "to promote the continued development of the Internet and other interactive computer services" and "to preserve the vibrant and competitive free market that presently exists for the Internet and other interactive computer services, unfettered by Federal or State regulation." American tech companies were thus allowed to innovate "with a minimum of government regulation." Digital governance, while not centralized or even coordinated in Washington, has emerged from a combination of other sources: the states, market forces, internal corporate codes of conduct, and the Federal Trade Commission's (FTC's) enforcement against unfair and deceptive business practices.

The deployment of sophisticated AI is challenging this global regulatory status-quo, much in the same way it is having a transformative effect on the technology sector. For example, while many experts in the field have wondered whether advances in AI are compatible with the onerous privacy dictates of the GDPR, given the enormous data sets required for training and machine learning, some astute observers, such as Theodore Christakis, ask about the other side of policy equation too: In the contest between the enormous benefits advanced AI offers society and the innovation-inhibiting impact of EU regulation, can the GDPR survive the allure of generative AI, notwithstanding the vast amount of data it demands?

The question of who—and what framework—will emerge as the world's leader in AI regulation is not clear, however, due to the uncertainty of exactly how advanced AI works, what it is capable of today, and what it will be able to do tomorrow. AI governance is further complicated because of the inescapably dual use of the technology—as hard as it will be to get a handle on commercial governance of AI, military applications will likely proceed on a separate track (potentially beyond conventional regulation). These kinds of open questions provide the United States with the opportunity, and perhaps the obligation, to assert leadership on global technology standards.

Of course, the European Union appears to have forged ahead by reaching an agreement on a comprehensive AI Act, which regulates the development, use, import, and distribution of high-risk and limited-risk AI systems. But even there, negotiations were disrupted by the pace of technological change. The original text of the AI Act, proposed in 2021, did not account for the rise of so-called generative AI and, in particular, those foundational or general purpose models such as OpenAI's GPT series that can serve as a platform for other applications. European industry leaders began to urge officials to think more carefully before adopting suffocating regulations that could hamstring entrants

in the sprint toward generative AI dominance. In any event, the European Parliament plenary vote on the final text of the EU AI Act will likely not occur until April. Moreover, as IAPP Research and Insights Director Joe Jones recently summarized, the act's final text indicates that the obligations for high-risk systems enumerated in Annex III of the act, including for use in biometrics, critical infrastructure, education, employment, essential private and public services, law enforcement, immigration, and the administration of democratic processes, will not apply until 24 months after the act comes into force. Obligations for those high-risk systems that form components of or are themselves products covered by the EU legislation listed in Annex II of the act, and are required to undergo conformity assessments before being placed on the market, will not apply until 36 months after the act comes into force. American agencies, by contrast, must meet a variety of substantive deadlines imposed by the Biden executive order before the end of 2024.

The United Kingdom, which, in contrast to the EU's characteristically precautionary and prescriptive posture, announced a pro-innovation approach to AI in March 2023, and also laid down a marker for global leadership by hosting an international AI Safety Summit at Bletchley Park in November 2023. The impetus for the U.K. summit was to focus world leaders' attention on the need for cooperation regarding the systemic risks that powerful AI could pose for global society at least as much as on the international competition to innovate. Building on the momentum of the Biden executive order, however, the United States did not take a back seat at the summit and instead exerted its influence over the direction of the talks. For example, Vice President Kamala Harris's remarks upon arriving in the U.K. for the Bletchley Park summit signaled the Biden administration's intent to lead on all dimensions of AI, including international policy. She noted that American domestic AI policies, in development even before generative AI, are intended to "serve as a model for global policy, understanding that AI developed in one nation can impact the lives and livelihoods of billions of people around the world."

It has become evident that either the United States has learned a lesson from its relative abstinence from global leadership on privacy policymaking, or AI is simply too important for America to stand on the international sidelines of Al governance. Indeed, work to promote the trustworthy use of Al within the federal government began as early as 2020 in response to the executive order on "promoting the use of trustworthy artificial intelligence in the federal government." The 2020 order catalyzed the impressive AI Risk Management Framework issued by the National Institute of Standards and Technology (NIST)—a framework that, like NIST's cybersecurity framework, is respected around the world. Additional actions by both the U.S. executive and legislative branches over the past year including the substantive 2023 Al Executive Order, fast-tracked congressional hearings driving toward the development of a bipartisan AI bill, and the establishment of the U.S. AI Safety Institute—signal the United States' active engagement on AI governance. These federal actions, in conjunction with frameworks developed by private companies such as Microsoft and Google, have set the path for U.S. global AI leadership. In a recent announcement regarding the creation of the U.S. AI Safety Institute Consortium, which includes over 200 leading AI stakeholders, key Biden administration members highlighted the importance of U.S. leadership. Secretary of Commerce Gina Raimondo remarked, "[t]hrough President Biden's landmark Executive Order, we will ensure America is at the front of the pack" when it comes to setting AI safety standards and protecting the AI innovation ecosystem. Bruce Reed, White House Deputy Chief of Staff, further commented that "[t]o keep pace with AI, we have to move fast and make sure everyone – from the government to the private sector to academia – is rowing in the same direction. Thanks to President Biden's landmark Executive Order, the Al Safety Consortium provides a critical forum for all of us to work together to seize the promise and manage the risks posed by AI."

U.S. global policy leadership on AI will likely play out based primarily on the Biden executive order issued just before the Bletchley Park summit, as well on the other key developments discussed below. What is also clear, though, is that this incipient, but muscular, U.S. leadership on AI governance is being conceived with what appears to be a multilateral mindset. This bodes well for the possibility of moving toward meaningful convergence on standards for the world's most powerful "safety impacting" and "rights impacting" AI systems.

The Biden Administration's Executive Order on Al

Issued on Oct. 30, 2023, the Biden administration's executive order on the "safe, secure, and trustworthy development and use of artificial intelligence" is one of the most significant federal efforts to regulate—or otherwise govern—information technology in modern history. Though it is not a self-contained, comprehensive piece of legislation such as the EU AI Act, the order has the immediate effect of directing federal agency resources toward assessing and addressing AI risk. The AI Act, by contrast, is not likely to come into full effect for the next several years.

The depth of President Biden's commitment to the project is reflected by the order's assignment of critical responsibilities for execution to some of the administration's most effective "go-to-to-get-it-done" administration officials, such as White House Deputy Chief of Staff for Policy Bruce Reed, who led the order's development, along with Secretary of Commerce Gina Raimondo. Secretary Raimondo, a Democratic policy "rock star," also played a key role in the development and execution of the CHIPS and Science Act, which among other things authorizes the use of federal funds to bolster domestic semiconductor production. She is now similarly tasked with implementing and overseeing certain key requirements in the AI executive order, including the unprecedented Defense Production Act (DPA) reporting and testing obligations imposed on the developers of the most powerful AI systems. Intense focus and coordination will be essential to carrying out a nearly overwhelming executive order that runs the gamut on AI-related risk, canvassing national security, cybersecurity, biosecurity, intellectual property, labor, antitrust, education, health care, privacy, consumer protection, and American leadership abroad.

Structure of the AI Executive Order

The executive order is different from the EU's draft AI Act, which sets forth a categorical approach to the regulation of certain AI use cases by the degree of risk they pose. In contrast, the order takes no position on which particular AI systems pose greater or lesser degrees of risk to the country's interests and values. It eschews "broad general bans or blocks" on the uses of generative AI in certain contexts, such as within the federal government itself. Instead, the document acknowledges the government's active responsibility to monitor and oversee the most powerful AI "foundation" or "frontier" models that could impact national security and critical infrastructure, and urges individual agencies to step in when AI might infringe on the rights and liberties of American citizens. In many respects, the executive order builds on early-stage federal AI frameworks—including the NIST AI Risk Management Framework 1.0, released in January 2023, and the White House AI Bill of Rights, published in October 2022 (both of which are referenced in the order)—and the voluntary commitments the president secured in July and September of last year from the country's leading AI companies to assure safety through various measures, including internal and external testing, information sharing and reporting, and watermarking.

Apart from the treatment of the highest computing capacity foundation models, the executive order outlines a largely sectoral, government-wide approach to the assessment and mitigation of Al-related risk, relying on agencies' existing legal authorities. It directs federal agencies to issue guidelines on an aggressive timeline. By October 2024, studies, guidelines, and reporting requirements are expected from myriad agencies, including the Department of Commerce, Health and Human Services, the Patent and Trademark Office, Treasury, Labor, Energy, Homeland Security, and Defense.

Significantly, a White House Artificial Intelligence Council, chaired by the White House deputy chief of staff for policy, was established to coordinate agency activity. Independent agencies, such as the FTC, were also encouraged to "exercise [] existing authorities," including "rulemaking authority," which, in the context of the FTC, includes actions to "ensure fair competition in the AI marketplace and to ensure that consumers and workers are protected from harms that may be enabled by the use of AI." On Nov. 21, 2023, the FTC adopted an omnibus resolution in an attempt to streamline the agency's ability to

issue civil investigative demands relating to "products and services that use or are produced using artificial intelligence."

Certain other agencies, such as the National Telecommunications and Information Administration within the Department of Commerce, have begun to set their own priorities in the wake of the order. In a speech delivered on Dec. 13, 2023, Assistant Secretary of Commerce for Communications and Information Alan Davidson outlined the agency's focus on open-source Al models, remarking that while a certain prominent venture capitalist recently opined that "[y]ou don't open source the Manhattan Project," the agency is nonetheless quite interested in reviewing both the risks and the benefits associated with publicly available model "weights." These "weights" are the parameters that determine how important any given variable is to a model's output, thereby playing a significant role in determining how Al systems make decisions.

Global Engagement and Leadership Under the Al Executive Order

The executive order explicitly calls for the United States to "lead the way" for global progress on AI and promote "responsible AI safety and security principles and actions with other nations, including our competitors." The White House fact sheet released with the order describes the administration's extensive global engagement already on AI governance, including with Japan regarding the Group of Seven's AI Code of Conduct, with the U.K. on its Safety Summit, and with India, the United Nations, and numerous others. Accordingly, the order expressly directs the secretary of state to engage in global leadership, including by "lead[ing] efforts to establish a strong international framework for managing the risks and harnessing the benefits of AI." And it likewise tasks the secretary of commerce to lead a coordinated initiative with key allies and international partners to develop consensus standards for AI.

Mandating Government Reporting and Red-Team Testing Requirements in the Interest of National Security

The AI executive order takes the position that the safe and trustworthy development of AI is a matter of national security. This posture is both true and legally relevant given that American constitutional law affords the executive considerable deference on matters of national defense and security. One of the most significant elements of the order, Section 4.2(a), directs the secretary of commerce to require "[c]ompanies developing or demonstrating an intent to develop potential dual-use foundation models to provide the Federal Government, on an ongoing basis, with information, reports, or records" relating to a series of topics, including training and development, ownership of model weights, and the results of mandatory red-team testing. Companies that acquire, develop, or possess large-scale computing clusters are obligated to report the existence and locations of such clusters and the total computing power available in each cluster.

The order sets forth initial technical standards to bring systems within scope of the Section 4.2(a) reporting requirement, namely those models "trained using a quantity of computing power great than 1026 integer or floating-point operations, or using primarily biological sequence data and using a quantity of computing power greater than 1023 integer or floating-point operations," and any "computing cluster ... physically co-located in a single datacenter," with a networking speed of over "100 Gbit/s" and with "a theoretical maximum computing capacity of 1020 integer or floating-point operations per second for training Al." It is worth noting that, per a series of revisions that postdate the Biden executive order, the threshold for high-impact general purpose models in the EU Al Act is now also staked to the level of computing power used for training—1025 floating-point operations, to be precise.

The order likewise imposes reporting requirements on such dual-use foundation models regarding physical and cybersecurity protections taken to ensure the integrity of model training activities and protection of model weights. And even before NIST develops guidance for conducting required redteam testing, the order itself mandates reporting of red-team test results regarding software

vulnerabilities, use of software to influence real or virtual events, and the possibility of AI self-replication or propagation, or lowering the barrier to entry for use of bioweapons by non-state actors. In the other words, the executive order is focused on assuring corporate responsibility and government awareness regarding existential threats.

To be sure, these standards are currently staked to only the highest thresholds of computing power and capacity. Some commentators suggested that no currently available models meet the requirements outlined by the order. However, as discussed, "developers of the most powerful AI systems" are already required to provide reports to the Department of Commerce. The White House further directs the secretary of commerce to define and update these standards on a regular basis.

The order sets forth the proposed statutory basis for the Section 4.2(a) reporting requirement, citing to the broadly powerful Defense Production Act (DPA). The fact sheet released by the White House confirms that the president's invocation of the DPA has compelled certain developers to report "vital information," including the results of safety testing, to the federal government. Originally enacted in 1950 in response to mobilization for the Korean War, the DPA affords the executive with certain emergency—and essentially unilateral—powers to regulate the economy in the interest of national security, including with respect to "unique technological requirements."

The DPA has been frequently amended and invoked to address modern-day challenges. In 1988, for example, Congress amended Section 721 of the act to provide a statutory basis for the Committee on Foreign Investment in the United States (CFIUS) review of foreign investments. Section 721 was further amended by the Foreign Investment and National Security Act of 2007 and the Foreign Investment Risk Review Modernization Act of 2018 to strengthen CFIUS review. And both the Trump administration and the Biden administration invoked the DPA with respect to the coronavirus to increase production of equipment deemed necessary for a response commensurate with the risk and impact of the global pandemic to American society.

Section 4517(b) of the DPA authorizes the president to take (or delegate the authority to take) actions to ensure that critical technologies are available from and restricted to reliable sources. Given that, under the executive order, "developers of the most powerful AI systems" are already required to submit reports to the Department of Commerce, Section 4.2(a) is self-executing, and no further action on the part of the secretary of commerce appears necessary. However, one would expect that the secretary will develop at least some ministerial reporting protocols and logistics, such as how and where to file reports and to whom reporting questions could be directed.

It remains to be seen whether there will be any serious industry resistance to AI reporting thresholds and testing requirements, and how federal courts will view any challenges. Nevertheless, if successful, the United States will have established one of the key elements for flexible and ultimately proinnovative AI regulation with the possibility, in reserve, to impose significant government intervention if and when it might prove necessary. In the meantime, the DPA route adopted in the order provides for considerable transparency between industry and government with respect to the development or acquisition, and the nature and behavior, of the most powerful, systemically significant AI technology.

Ideally, this reporting will be more informative to policymakers than it is burdensome for industry participants, or stifling of innovation that would benefit society. The reporting and testing would allow American officials to protect domestic interests, yes, but also to advocate more effectively for appropriate global standards.

The importance of global cooperation on AI standards is evident when one considers other actions taken by the Department of Commerce under the executive order. To implement Section 4.2(c), and to address the risks associated with the ability of malicious foreign actors to "engage in certain AI training runs" on U.S. "large-scale computing infrastructure," the Department of Commerce recently issued a Notice of Proposed Rulemaking that would impose government reporting requirements on all

U.S. providers of Infrastructure-as-a-Service (IaaS) products. These requirements include, amongst other things, providing notice to the department whenever the IaaS provider has "knowledge" of a "transaction by, for, or on behalf of a foreign person which results or could result in the training of a large AI model with potential capabilities that could be used in malicious cyber-enabled activity." Alongside providing useful information to the U.S. government, these requirements may pose challenges for U.S. infrastructure providers regarding how to report such information in a manner that respects existing privacy commitments and relevant law. They may generate frustration amongst allied governments if information is collected on domestic champions without further coordination. It may be the case that the U.S. government will need to engage in discussion with partner governments and others to develop a more durable agreement.

Congressional Action?

In the fact sheet accompanying the executive order, the White House also speaks of working with Congress on bipartisan AI legislation so that the United States can lead global responsible innovation for the technology. The order's urgency and accelerated timelines are likely requirements that reflect not only the speed with which the underlying technology is advancing but also the impending electoral season—with its possibility of executive turnover and the certainty of political distraction throughout the Capitol.

In his remarks announcing the executive order, President Biden referenced existing congressional initiatives, such as Senate Majority Leader Chuck Schumer's (D-N.Y.) SAFE Innovation Framework, that are building toward the development of bipartisan AI legislation. A cornerstone of Schumer's approach is to fast-track the development of such legislation by bringing together members of Congress, industry leaders, and members of civil society in a series of AI Insight Forums. Those forums have concluded, and reports indicate that legislative drafting has commenced. However, some senators have expressed frustration with the nature of the process and have moved forward with their own AI legislation. Sen. Josh Hawley (R-Mo.), for example, long interested in opening up avenues for plaintiffs to sue technology companies directly, recently attempted to introduce a bill to clarify that Section 230's liability shield does not apply to generative AI. Whether a truly bipartisan agreement on AI legislation can be reached in Congress in an election year is by no means assured.

U.S. Leadership at the Bletchley Al Safety Summit

Two days after the release of the Biden AI executive order, Vice President Harris traveled to the U.K. to attend the AI Safety Summit. On the eve of the summit, she delivered a speech at the American Embassy in London to outline American leadership on the issue, including the capacity of the United States to address the concrete and present risks of AI. The vice president also took the opportunity to announce the joinder of 30 countries (not including China, as of Jan. 12) to the U.S.-led "Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy," and the launch of the U.S. AI Safety Institute (AISI), a body similar to the U.K. AI Safety Institute that will create standards to test the safety of AI models for public use. On Feb. 7, Secretary Raimondo named Elizabeth Kelly, currently Special Special Assistant to the President for Economic Policy at the White House National Economic Council, as AISI's inaugural director. Japan has also signaled an intention to create a safety institute, and the AI Office established by the European Commission may serve a similar function. These safety institutes may emerge as key focal points for collaboration around the development and conducting of evaluations of highly capable AI models across governments, and a key plank of global coordination on regulation.

The Bletchley summit appeared to be a diplomatic success, given that participating countries—which included not only leading democratic nations but also, most notably, China—signed the first international declaration to cooperate on Al safety. The signatories agreed that "frontier" Al models, or those "highly capable general-purpose Al models, including foundation models," pose particular safety risks. They agreed to "support an internationally inclusive network of scientific research on frontier Al

safety that encompasses and complements existing and new multilateral, plurilateral and bilateral collaboration." The participants also committed to further meetings in Korea and France in the next 12 months.

Nevertheless, one of the key elements of the declaration is transparency. In the context of an internationalist approach, transparency rests on voluntary commitments from private firms. As discussed, previous actions by the Biden administration had already secured such voluntary commitments from leading AI developers. The executive order goes one step further by directing the secretary of commerce to require government reporting requirements. The United States has therefore positioned itself as willing to exercise the "hard" power of the DPA to achieve leadership in the field of AI regulation that might otherwise be limited to so-called soft power.

OMB Implementing Guidance for Federal Agencies, and Requirements for "Safety-Impacting" or "Rights-Impacting" AI

Directly following the issuance of the AI executive order, the Office of Management and Budget (OMB) released a draft policy titled "Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence." If adopted, this guidance would promulgate new guidelines for the use of AI in government in line with existing requirements such as the Artificial Intelligence in Government Act of 2020, the Advancing American AI Act, and President Biden's AI executive order. It would establish AI governance structures in all federal agencies. What's more, implementing the Biden AI order, agencies must appoint a chief AI officer to promote innovative uses of AI and manage risk. Agencies would also be required to publish their strategies for AI use, including by outlining each agency's current and planned top use cases for AI models.

Interestingly, and perhaps emulating the EU's approach in the draft AI Act, the OMB guidance takes a categorical approach to AI regulation, establishing certain minimum requirements for uses of AI that are presumed to be "safety-impacting" or "rights-impacting." Examples of safety-impacting AI outlined by OMB include AI that is used to control or influence the functioning of critical infrastructure, delivery of biological or chemical agents, and the movement of vehicles; examples of rights-impacting AI include use of AI in decisions that impact protected speech, law enforcement, immigration, loan allocation, and child welfare.

Uses of AI that fall within the specified categories must follow certain minimum practices. These include ensuring periodic human review of AI systems; notifying individuals when AI meaningfully influences the outcome of decisions that specifically impact them, such as the denial of benefits; and completing AI impact assessments. AI impact assessments under the new OMB guidance must document the intended purpose for the AI and its expected benefit; the potential risks of using AI; and the quality and appropriateness of the data used in the design, development, training, and testing of the AI.

The OMB guidelines contemplate a future in which AI is heavily integrated into federal agency decision-making. This raises a variety of interesting questions, including how courts will review agency decisions that are AI based. The robustness of these draft guidelines, which implement the AI executive order's directives, signal an acute level of intensity on the part of federal government officials to assess and address AI risk. The guidance could also serve as a model for private companies that are beginning to build their own internal AI governance programs.

Overall, the impact of the U.S. government's leadership on AI so far—with the executive order, OMB guidelines, the White House AI Bill of Rights, NIST's AI Risk Management Framework, and voluntary commitments by leading AI companies—is world leading. The U.S. approach remains sectoral, particularly in comparison to comprehensive legislation such as the prospective EU AI Act. However,

the depth of analysis and coordination on the part of American agencies that has already begun and will continue to unfold based on presidential executive orders is staggering. And it is perhaps this expanded and aggressive sectoral approach that will allow for the development of regulation that is nimble enough to adapt to and address the technology itself, avoiding the pitfalls of processes such as the EU AI Act negotiations, which were disrupted by the unexpected launch of frontier models. At the same time, as noted at the outset of this piece, the intensity of the new order has instigated a backlash among some who view its regulatory mandates as overreaching, unauthorized, and unwarranted.

Nevertheless, the impacts of these governance elements in the United States will be widespread in both government and civil society. And with the specter of the Defense Production Act over the development of the most powerful AI systems, there may be an effective fail-safe at the ready. These American AI policy initiatives are certainly substantive and impressive, and they could well serve as models for our international partners and for the private sector at home.