Generative AI:
The Future of
Innovation Power

SPECIAL EDITION REPORT

SOCIETY PANEL

Governance of Generative AI

SPECIAL COMPETITIVE STUDIES PROJECT

Artificial Intelligence Primer

With the emergence and rapid adoption of generative artificial intelligence (GenAI), the landscape of our world significantly shifted over the last year. The current capabilities and future potential of GenAI are now at the very heart of the national conversation.

GenAI is a category of algorithms that finds patterns in training datasets and makes generalizations to generate content such as text, images, or audio, given natural language or multimedia input. Individuals can interact directly with GenAI tools via natural language interface chatbots like ChatGPT or Bard, or through application portal interfaces (APIs) that connect software systems. In earlier paradigms, humans gave computers precise commands in software languages. In the emerging paradigm, humans provide computers with intent in natural language, and the computer determines how to execute it. [1]

Foundation models like GPT-4 and PaLM 2 have increased in size and capability at an incredible rate, moving in a short time from curiosities to tools that match or exceed human performance in a number of areas.[2] The rate of adoption has been equally impressive, driven in part by the low entry barrier to GenAI. Organizations or individual end-users that have not developed the robust datasets associated with other types of machine learning can still use foundation models' generalized capabilities, or fine-tune a model for a specific purpose. As a result, millions of users can generate text, images, code, and other outputs, often with much less skill required than before.[3] These capabilities will increasingly drive innovation across many sectors, from healthcare to defense.

1 Jakob Nielsen, AI: First New UI Paradigm in 60 Years, Nielsen Norman Group (2023).
2 Dmytro Nikolaiev, Behind the Millions: Estimating the Scale of Large Language Models, Medium (2023); Reed Albergotti, The Secret History of Elon Musk, Sam Altman, and OpenAI, Semafor (2023); Helen Toner, The Illusion of China's AI Prowess, Foreign Affairs (2023); GPT-4 Technical Report, Open AI (2023).
3 Krystal Hu, ChatGPT Sets Record for Fastest-Growing User Base – Analyst Note, Reuters (2023).

GenAI has limitations. It can infer incorrect information, or draw on training data that is itself nonsensical, or contains incorrect information.[4] However, it is very likely that model outputs will improve significantly in the coming months and years.[5] For now, as humans interact with LLMs, they will need to develop ways of recognizing model limitations while still producing useful outputs.

GenAI opens a range of new capabilities and applications that will shape the geopolitical balance, changing the calculus for military, diplomatic, and economic power, as well as societal cohesion.

Even at this early stage, it is apparent that GenAI opens a range of new capabilities and applications that will shape the geopolitical balance, changing the calculus for military, diplomatic, and economic power, as well as societal cohesion. What's more, GenAI will likely exacerbate existing threats such as disinformation, cybersecurity, and biorisks by way of its diffusion. It also may create novel risk categories, such as algorithmic misalignment, with which we are only beginning to grapple.

The national security implications for our nation are profound and the world is rushing to make sense of shifting opportunities and threats. Governments, industry, and academia will need to develop adaptive strategies capable of responding to more changes in an already turbulent environment. It is paramount to underscore the role the U.S. government needs to play at this critical juncture. The government must become an active participant in shaping the future if it is to be ready, informed, and organized to face the strategic implications of GenAI. We cannot afford endless debate and hype-driven paralysis.

This memorandum to the President of the United States and Congress presents the near-term implications of Governance of GenAI, along with policy recommendations to adapt to rapidly changing conditions. We believe that with careful planning and ethical considerations, the United States can harness GenAI to secure a future where AI serves the nation, bolsters our defenses, and enhances our collective security.

4 Frank Neugebauer, Understanding LLM Hallucinations, Medium (2023).
5 Sarah Wang & Shangda Xu, The Next Token of Progress: 4 Unlocks on the Generative AI Horizon, Andreessen Horwitz (2023).

Governance Memo

MEMORANDUM TO THE PRESIDENT OF THE UNITED STATES AND CONGRESS

FROM: Special Competitive Studies Project

SUBJECT: Governance of Generative AI (GenAI)

Two different timeline trajectories are shaping the current GenAI governance landscape. The development of GenAI is evolving rapidly, while GenAI governance mechanisms[6] are moving slowly.

We have seen OpenAI's release of GPT-4 in March 2023, with significantly advanced capabilities compared to GPT-3.5, released in November 2022.[7] At the same time as frontier GenAIs are advancing, we have seen a proliferation of open-source models.[8] Open-source models democratize access to GenAI capabilities. While increased access has positive domestic implications (e.g., increased domestic competition as opposed to limiting the market to a few), it also means AI capabilities are available to U.S. adversaries and malign non-state actors who could potentially use them to harm us. This environment is still developing, but as open-source capabilities advance, the United States will have a clearer picture of the harms that can be advanced against U.S. interests from open-source GenAI systems.

In contrast, regulation rightly moves slowly in the United States. Governing GenAI to align with democratic values will take time. Employing regulation that harnesses opportunities from GenAI and mitigates its harms requires a well-informed picture of GenAI's implications and threats. Effective U.S. GenAI regulation requires aligning with both of these realities.

The United States must leverage and expand existing governance authorities by utilizing GenAI tools and upskilling regulators with the necessary expertise. The United States must take the following specific actions, in order of urgency: (1) Protect our digital information

6 "Governance" includes regulation as well as non-regulatory mechanisms (e.g., self-governance, independent auditing, advocacy, philanthropy).

7 Jon Martindale, GPT-4 vs. GPT-3.5: How Much Difference Is There?, Digital Trends (2023).

8 Davide Castelvecchi, Open-Source AI Chatbots Are Booming — What Does This Mean for Researchers?, Nature (2023).

The United States must leverage and expand existing governance authorities by utilizing GenAI tools and upskilling regulators with the necessary expertise.

and elections systems by convening stakeholders to agree to a synthetic media code of conduct for elections, passing legislation to assign a lead agency for alerting the public of synthetic media use in federal elections, and encouraging department and agency heads to use all available regulatory tools to adopt proposed public digital literacy education and disinformation awareness ahead of 2024 U.S. elections; (2) Find ways to help regulators identify AI uses cases that have highly consequential impacts on society based on their sector-specific contexts so that they can focus regulatory efforts on GenAI that has significant beneficial outcomes to society while mitigating the worst of the harms;[9] (3) Address threats posed by foreign digital platforms from countries of concern by tailoring restrictions to specific platforms like TikTok and subsequently establishing a comprehensive risk-based policy framework; (4) Over time consider establishing a centralized AI authority that can regulate AI issues that cut across sectors and fill regulatory gaps in sectors;[10] and (5) Establish under the G20 an international Forum on AI Risk and Resilience (FAIRR) that convenes key states and private actors to build a governance floor for managing GenAI tools' malign non-state use, potential state-based infringements on other states' sovereignty, and injurious societal impacts.

PURPOSE

GenAI is a Top Priority for Governance

The United States has long-established robust governance mechanisms that can be leveraged to address imminent GenAI threats to our society. At the same time, the United States must explore new approaches and authorities to address needs that are not met by existing authorities. In addition, given that GenAI crosses borders and jurisdictions, the United States and other nations need to explore international governance mechanisms for addressing global issues raised by GenAI.

9 To balance the need for regulation without stifling innovation, the United States must focus regulation on AI that will have highly consequential beneficial or harmful impacts on society.

10 Immediate actions to govern GenAI outcomes must be taken in parallel with exploring the longer term goal of potentially establishing a new AI authority.

NEAR TERM IMPLICATIONS

GenAI's Impact on Society

Near-term implications of GenAI result from its capabilities and adoption pace. In the next 12 to 18 months, GenAI's acceleration of disinformation presents an imminent and pressing threat to trust in democratic institutions domestically and globally.[11] GenAI capabilities — including the creation of synthetic media,[12] such as deepfake images and synthetic audio — have advanced on the cusp of the 2024 election cycle, during which well over one billion people worldwide[13] will go to the polls.[14] While disinformation challenges to elections are nothing new, synthetic media adds qualitative and quantitative elements previously not present. Synthetic media is nearly, or at least soon will be, indiscernible from truth and GenAI allows for increased volume. These capabilities introduce heightened national security vulnerabilities by providing new attack surfaces to adversaries and malign non-state actors. GenAI also increases the risks posed by content distribution platforms – particularly those with the potential to come under the direct or indirect control of foreign governments in countries of concern – by making it easier to collect private data given increased platform engagement, influence users, and polarize society.[15]

Near-term implications of GenAI result from its capabilities and adoption pace.

11 Tiffany Hsu & Steven Lee Meyers, A.I.'s Use in Elections Sets Off a Scramble for Guardrails, New York Times (2023).

12 "[S]ynthetic media, also referred to as generative media, is defined as visual, auditory, or multimodal content that has been artificially generated or modified (commonly via artificial intelligence). Such outputs are often highly realistic, would not be identifiable as synthetic to the average person, and may simulate artifacts, persons, or events." See Responsible Practices for Synthetic Media: A Framework for Collective Action, Partnership on AI at 3 (2023).

13 The 2024 election calendar includes elections not only in the United States, but also in Taiwan, Indonesia, South Korea, India, the European Union, Mexico, Egypt, South Africa, and more.

14 See Mekela Panditharatne & Noah Giansiracusa, How AI Puts Elections at Risk — And the Needed Safeguards, Brennan Center for Justice (2023); Thor Benson, Brace Yourself for the 2024 Deepfake Election, Wired (2023).

15 The risks associated with TikTok and other foreign digital platforms from countries of concern are anticipated to grow with GenAI. See Meaghan Waff, TikTok Is the Tip of the Iceberg: National Security Implications of PRC-Based Platforms, Special Competitive Studies Project (2023).

The rapid advancement and adoption of GenAI applications[16] make addressing AI threats to democratic values such as privacy, non-discrimination, fairness, and accountability, more urgent.

GenAI's Impact on Global Governance

Governing GenAI in a manner consistent with democratic values is, first and foremost, a domestic imperative. Regardless of international conditions, the U.S. government has an obligation to ensure that innovation's impacts on society accord with U.S. constitutional rights and the democratic processes that support those rights. Yet, that fundamental reality does not obscure that the GenAI revolution is occurring amid a significant international shift toward an ideological contest of governance models.[17]

The U.S. government has an obligation to ensure that innovation's impacts on society accord with U.S. constitutional rights and the democratic processes that support those rights.

Around the world, AI governance models are on display for judgment and, ultimately, adoption.[18] The European Union has put forward its archetype in the EU AI Act.[19] Likewise, the People's Republic of China (PRC) is moving to regulate GenAI.[20] How these governance regimes either spur or stifle innovation and economic opportunity while furthering or curtailing human rights will impact the attractiveness of their approaches around the globe.[21] The United States should not underestimate these alternatives — including the PRC's. On paper, at least, Beijing's recent rules offer data privacy protections, worker protections,

16 Krystal Hu, ChatGPT Sets Record for Fastest-Growing User Base, Reuters (2023). Stability AI's Stable Diffusion hit 10 million daily users. See Mureji Fatunde & Crystal Tse, Stability AI Raises Seed Round at $1 Billion Value, Bloomberg (2022). Recent trends in GenAI include plugins connecting GenAI models and third party data. See Jason Nelson, These ChatGPT Plugins Can Boost Your Productivity With AI, Yahoo! Finance (2023); Kyle Wiggers, Microsoft Goes All in on Plug-ins for AI Apps, TechCrunch (2023).

17 Mid-Decade Challenges to National Competitiveness, Special Competitive Studies Project at 16-27 (2022).

18 See Anu Bradford, The Race to Regulate Artificial Intelligence, Foreign Affairs (2023).

19 EU AI Act: First Regulation on Artificial Intelligence, European Parliament News (2023).

20 Qianer Liu, China to Lay Down AI Rules with Emphasis on Content Control, Financial Times (2023).

21 Whether the PRC can spur innovation with regulations designed to ensure the Chinese Communist Party's ultimate control remains to be seen. Qianer Liu, China to Lay Down AI Rules with Emphasis on Content Control, Financial Times (2023); Matt Sheehan, China's AI Regulations and How They Get Made, Carnegie Endowment for International Peace (2023).

and seek to prevent discrimination.[22] The appeal of the American experiment — its ability to foster innovation and harness significant benefits for society, while respecting individual rights and protecting society from the worst of the harms — is under scrutiny. Providing an effective democratic GenAI governance model that other democratic nations adopt will shape the future geopolitical order.

Providing an effective democratic GenAI governance model that other democratic nations adopt will shape the future geopolitical order.

U.S. GOVERNMENT APPROACH TO GENERATIVE AI: ACTING WHILE LEARNING

U.S. government attention to AI has grown exponentially in recent years. Congress has passed new legislation[23] and the Executive Branch[24] has prioritized action on internal government adoption, broader responsible development and use, and ensuring U.S. AI leadership globally. In recent months, the government has turned its attention to GenAI, seeking outside expertise on how to address the novel risks and opportunities presented by the technology.[25] The White House announced a "voluntary commitment" from the seven

22 Qianer Liu, China to Lay Down AI Rules with Emphasis on Content Control, Financial Times (2023).

23 See, for example, Pub. L. 116-283, National Artificial Intelligence Initiative Act of 2020 Div. E (2021); Pub. L. 116-260, AI in Government Act of 2020, Div. U (2020); Pub. L. 117-167, CHIPS and Science Act (2022); and Pub. L. 116-258, Identifying Outputs of Generative Adversarial Networks Act (2020).

24 Examples of Executive Branch actions include: the work of the National AI Initiative Office to coordinate AI policy; the Office of Science and Technology Policy (OSTP) releasing a Blueprint for an AI Bill of Rights and the National Institute of Standards and Technology (NIST) releasing an AI Risk Management Framework – both aiming to provide guidance on the responsible development and use of AI; Federal regulatory agencies announcing in April their shared commitment to mitigate bias and discrimination through application of their existing authorities to AI systems; export control regulators have moved to curb competitor's access to chips critical to powering AI; and issuing Executive Orders. See Legislation and Executive Orders. See National AI Initiative Office (last accessed 2023); Blueprint for an AI Bill of Rights, The White House (last accessed 2023); AI Risk Management Framework, National Institute of Standards and Technology (last accessed 2023); Justice Department's Civil Rights Division Joins Officials from CFPB, EEOC and FTC Pledging to Confront Bias and Discrimination in Artificial Intelligence, U.S. Department of Justice (2023); Commerce Implements New Export Controls on Advanced Computing and Semiconductor Manufacturing Items to the People's Republic of China (PRC), Bureau of Industry and Security, Department of Commerce (2022); EO 13859, Maintaining American Leadership in Artificial Intelligence (2019); EO 13960, Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government (2020).

25 One example includes U.S. Copyright Office listening sessions on how to address the intersection of generative AI tools and copyrighted materials and the use of copyrighted materials to train generative AI tools. Copyright and Artificial Intelligence, U.S. Copyright Office (last accessed 2023). See also AI Inventorship Listening Session - East Coast, U.S. Patent and Trademark Office (2023).

As the GenAI revolution sweeps the nation and the world, it will continue to impact the U.S. government's governance mechanisms as much as the substance of what is governed.

foremost GenAI companies to ensure safety, security, and trust in their models prior to public release.[26] Separately, the White House is seeking public input as it builds a National AI Strategy,[27] and the President's Council of Advisors on Science and Technology established a GenAI working group to advise the White House as it develops GenAI policy.[28] Additionally, the National Institute of Standards and Technology (NIST) announced the establishment of a public, collaborative Generative AI Working Group.[29] In Congress, GenAI educational briefings[30] and hearings[31] populate the calendar, and a flurry of AI legislative and framework proposals have been introduced.[32]

26 See Fact Sheet: Biden-Harris Administration Secures Voluntary Commitments from Leading Artificial Intelligence Companies to Manage the Risks Posed by AI, The White House (2023). Following these commitments, four of the technology companies launched the Frontier Model Forum, a collaboration to develop frontier AI models safely and responsibly. Rebecca Klar, Top Tech Companies Create Joint AI Safety Forum, The Hill (2023).

27 The White House's Office of Science and Technology Policy issued a request for information on how the U.S. government should approach various aspects of AI, including how to incorporate GenAI into operations, and whether laws and policies may need to be updated to account for AI. See 88 Fed. Reg. 34194, Request for Information; National Priorities for Artificial Intelligence, The White House, Office of Science and Technology Policy (2023).

28 See PCAST Working Group on Generative AI Invites Public Input, The White House (2023).

29 The NIST Generative AI Working Group seeks to "address the opportunities and challenges associated with AI that can generate content, such as code, text, images, videos and music." Biden-Harris Administration Announces New NIST Public Working Group on AI, U.S. National Institute of Standards and Technology (2023).

30 Senators Chuck Schumer (D-NY), Martin Heinrich (D-NM), Mike Rounds (R-SD), and Todd Young (R-IN) announced in a June 2023 Dear Colleague letter that they are spearheading a series of AI educational briefings to all Senators. See Leader Schumer Leads Bipartisan Dear Colleague Letter – With Senators Rounds, Heinrich, And Young – Announcing Three Bipartisan Senators-Only Briefings This Summer, Including First-Ever Classified All-Senators AI Briefing, Senate Democrats (2023).

31 Congressional hearings have covered generative AI impacts to human rights, intellectual property, Department of Defense operations, and governance. See Artificial Intelligence and Human Rights, U.S. Senate Committee on the Judiciary (2023); Artificial Intelligence and Intellectual Property – Part I: Patents, Innovation, and Competition, U.S. Senate Committee on the Judiciary (2023); Hearing to Receive Testimony on the State of Artificial Intelligence and Machine Learning Applications to Improve Department of Defense Operations, U.S. Senate Committee on Armed Services (2023); Oversight of A.I.: Rules for Artificial Intelligence, U.S. Senate Committee on the Judiciary (2023).

32 For example, with an eye toward the impacts of generative AI, in June 2023, Senate Majority Leader Chuck Schumer announced a SAFE Innovation Framework, which outlines policy objectives for governing AI while fostering continued innovation. See Majority Leader Schumer Delivers Remarks To Launch SAFE Innovation Framework For Artificial Intelligence At CSIS, Senate Democrats (2023). Senators Josh Hawley and Richard Blumenthal held a hearing on July 26, 2023 to discuss guiding principles for regulating AI. See Hawley, Blumenthal Hold Hearing On Principles For Regulating Artificial Intelligence, Senators Josh Hawley and Blumenthal (2023).

As the GenAI revolution sweeps the nation and the world, it will continue to impact the U.S. government's governance mechanisms as much as the substance of what is governed. To address these impacts, the U.S. government requires the increased capacity to both adopt[33] and govern GenAI tools. While government collaboration with external GenAI experts and stakeholders is paramount, policymakers and regulators fulfilling their mandates will require expanded in-house capacity and expertise in data science and AI in order to understand these systems' impacts on their area of focus.[34] Regulators, in particular, will need new capabilities to investigate and take action where appropriate. Importantly, the U.S. government should utilize existing capabilities[35] to address the most pressing concerns presented by GenAI while continuing to explore additional mechanisms.

---

33 GenAI presents regulators with new tools that can be integrated into their existing sectoral operations. For example, large language models (LLMs) will provide regulators greater ability to query existing governmental databases to better inform policy making with long-term records and trend lines. GenAI's ability to draft text and manage administrative processes will free bandwidth for officials to spend more time performing higher-level cognitive tasks. Simultaneously, for both regulators and regulated entities, GenAI tools will improve information sharing between them. GenAI will assist the former in compliance and enforcement and the latter in clarity in determining applicable regulations and compliance. Thus, GenAI tools hold the potential to allow regulators and regulated entities to invest greater time and energy in smart solutions and decrease regulatory compliance burdens. See Applied AI Challenge: Large Language Models (LLMs), General Services Administration (2023).

34 To satisfy the need for more expertise, the U.S. government requires more pathways for recruiting and retaining technology talent in government, increased public-private partnerships opportunities for research and development and governance, greater resources for digital infrastructure, and expanded contracting flexibility. See Final Report, National Security Commission of Artificial Intelligence (2021); Mid-Decade Challenges to National Competitiveness, Special Competitive Studies Project (2022).

35 Adapting available capacities to better govern the GenAI space depends, at first principles level, on the existence of capable authorities for federal regulators. This memo would be remiss if it did not mention the growing foundational challenge of judicial developments, particularly surrounding the tension between the doctrines of Chevron Deference and Major Questions. These judicially applied doctrines concern, respectively, federal courts giving deference to decisions by administrative agencies and the ability of Congress to provide broader grants of decision-making authority to those agencies. Recent judicial developments have begun to weaken and create uncertainty for regulatory powers involving economic national security tools and domestic regulation that intersects with foreign policy matters by shifting policy authority to the courts, which might not be as well versed in the domain. The application of these doctrines could reverberate across the federal government with severe impacts on national security. See Walter Johnson & Lucille Tournas, The Major Questions Doctrine and the Threat to Regulating Emerging Technologies, Santa Clara High Technology Law Journal (2023); Amy Howe, Supreme Court Will Consider Major Case on Power of Federal Regulatory Agencies, SCOTUS Blog (2023); Niina H. Farah & Lesley Clark, Supreme Court Axes Debt Relief, Threatens Climate Regs, E&E News (2023). Executive and Congressional actions in response could reduce resulting uncertainty or prepare the space for a new regulatory environment. In the Executive Branch, the Department of Justice should increasingly raise the Court's awareness that these decisions that seem remote from them are actually closely tied to U.S. national security. See Timothy Meyes & Ganesh Sitaraman, The National Security Consequences of the Major Questions Doctrine, Michigan Law Review (2023). Simultaneously, the Congress should consider this doctrinal trend when legislating and take steps to ensure it has the capability to draft more precise and nuanced statutory law, of the type often left to regulations, should regulators lose those powers. Maya Kornberg & Martha Kinsella, Whether the Supreme Court Rolls Back Agency Authority, Congress Needs More Expert Capabilities, Brennan Center for Justice (2023).

WAY FORWARD

Recommended Actions

Domestic Election Systems. Protecting U.S. digital information and elections systems requires three simultaneous actions:

1. NIST should convene industry to agree to a voluntary standard of conduct for synthetic media around elections in advance of the 2024 U.S. elections.[36] Industry should use existing ethical guidance, such as the Partnership on AI's Responsible Practices for Synthetic Media,[37] to inform the new code of conduct.

2. Congressional leaders should work to scale public digital literacy education and disinformation awareness by: (1) passing legislation to assign a lead agency to alert the public of synthetic media use in federal elections, and (2) encouraging department and agency heads to use all available regulatory tools to build public resilience against disinformation under the guidance of the lead agency.

Congressional leaders should work to scale public digital literacy education and disinformation awareness.

   a. First, Congress should authorize a federal entity (e.g., the Department of Homeland Security in coordination with expert agencies such as the Federal Election Commission, NIST,[38]

36 The recent White House convening of leading industry GenAI actors which resulted in voluntary commitments can serve as a model for convening stakeholders such as content distribution platforms. Fact Sheet: Biden-Harris Administration Secures Voluntary Commitments from Leading Artificial Intelligence Companies to Manage the Risks Posed by AI, The White House (2023).

37 See Responsible Practices for Synthetic Media: A Framework for Collective Action, Partnership on AI (2023).

38 These agencies would have jurisdiction over domestic identification of synthetic media use, whereas the Intelligence Community has jurisdiction over foreign actors' use of synthetic media.

and relevant Intelligence Community partners[39] as necessary) to take charge of documenting and alerting the public to the use of synthetic media in federal elections, assessing authenticity and attribution in highly consequential use cases, and taking proactive steps to increase public digital literacy in elections. Without overarching federal legislation, the U.S. government risks a patchwork of state-level regulatory frameworks.[40]

b. Second, Congress should encourage governmental entities with existing counter-disinformation and election integrity efforts to adopt measures to scale public digital literacy education and disinformation awareness ahead of the 2024 elections.[41] Relatedly, the U.S. government should encourage the private sector to continue to collaborate to identify potential authentic or inauthentic networks impacted by AI-enabled disinformation and synthetic media. Congress should also enact legislation clarifying the Federal Election Commission's authorities to regulate the use of deepfakes in federal elections.[42]

3. To encourage transparency and increase safety, content distribution platforms[43] should be required to technically support a content and provenance standard, such

39 The U.S. Intelligence Community has a leading role in combating foreign malign influence. Foreign malign influence is defined as "any hostile effort undertaken by, at the direction of, or on behalf of or with the substantial support of, the government of a covered foreign country with the objective of influencing, through overt or covert means." See 50 U.S.C. §3059, Foreign Malign Influence Center. Examples of existing efforts include the Office of the Director of National Intelligence's Foreign Malign Influence Center and the Federal Bureau of Investigation's Foreign Influence Task Force. Related government efforts also include the Cybersecurity and Infrastructure Security Agency "MDM Team" and Department of State's Global Engagement Center. Knowledge sharing with domestic entities around how to identify deepfakes and other synthetic content would further protect against synthetic media use for elections. For example, Cyber Command and NSA could share lessons learned and capabilities from how their experience protecting elections from foreign influence transfers to synthetic media. See How NSA, U.S. Cyber Command are Defending Midterm Elections: One Team, One Fight, National Security Agency/Central Security Service (2022).

40 Sixteen states have introduced or enacted legislation to restrict the use of deepfakes. These include: California, Connecticut, Delaware, Georgia, Hawaii, Illinois, Louisiana, Massachusetts, Minnesota, New Jersey, New York, Rhode Island, Texas, Virginia, Washington, Wyoming. See Isaiah Poritz, States Are Rushing to Regulate Deepfakes as AI Goes Mainstream, Bloomberg (2023).

41 These interventions include, but are not limited to pre-emptive education efforts for the public such as digital literacy. See generally: Emily K. Vagra & Melissa Tully, News Literacy, Social Media Behaviors, and Skepticism Toward Information on Social Media, Information, Communication & Society (2019); Andrew Guess, et al., A Digital Media Literacy Intervention Increases Discernment Between Mainstream and False News in the United States and India, PNAS (2022).

42 See Karl Evers-Hillstrom, AI-Generated Deepfake Campaign Ads Escape Regulators' Scrutiny, Bloomberg Law (2023).

43 The information environment surrounding elections is a prime concern, but only one of many, with respect to the safety of content distribution platforms. In our democratic society, it is difficult to draw bright lines between content directly pertaining to an election and the broader information space that informs individuals' political and economic decisions.

as the Coalition for Content Provenance and Authenticity technical standards, that identifies whether content is GenAI generated or modified.[44] A trusted federal entity, to be determined by Congress, should monitor the platforms and enforce these transparency levers as circumstances befit, while safeguarding the platforms' intellectual property.

Domestic Regulatory Needs. The United States should consider a flexible AI governance model, which would cover GenAI, in accordance with four key principles previously identified by SCSP:[45] (1) Govern AI use cases and outcomes by sector; (2) Empower and modernize existing regulators, while considering a longer-term centralized AI regulatory authority that can address gaps as well as sector-cross-cutting issues; (3) Focus on highly consequential uses, both beneficial and harmful significant impacts;[46] and (4) Strengthen non-regulatory AI governance, such as the voluntary codes of conduct, with input from industry and key stakeholders.

The United States should consider a flexible AI governance model, which would cover GenAI, in accordance with four key principles previously identified by SCSP.

The United States has existing, robust regulatory mechanisms that can be employed to address concerns raised by GenAI use. Given that GenAI opportunities and challenges are inextricably tied to the contexts in which it is used, the United States should continue adapting present sector-specific regulatory authorities to address issues raised by GenAI adoption. Regulatory bodies should apply their existing authorities to GenAI and be empowered with the necessary skills and expertise.[47] Congress should also legislate requirements that operationalize responsible and ethical AI principles. For example, legally requiring industry

44 C2PA Specifications, Coalition for Content Provenance and Authenticity (2023).
45 Mid-Decade Challenges to National Competitiveness, Special Competitive Studies Project at 87 (2022).
46 This shares the risk-based approach common to both the EU's AI Act and the NIST AI Risk Management Framework. See Regulatory Framework Proposal on Artificial Intelligence, European Commission (2023).
47 On the application of existing legal authorities to "automated systems," including AI, see Joint Statement on Enforcement Efforts Against Discrimination and Bias in Automated Systems, U.S. Consumer Financial Protection Bureau, Department of Justice, Equal Employment Opportunity Commission, and Federal Trade Commission (2023).

to provide information about the data used to train commercial GenAI[48] and the model itself would operationalize "transparency," a common responsible and ethical AI principle.[49]

Regulators will not be able to regulate every AI model or tool — nor should they have to. To balance enabling AI innovation against regulation, regulators should expend their oversight efforts on AI use cases that are highly consequential to society. Specifically, regulators should focus on encouraging AI that has significant benefits and mitigating the worst of the harms. To do this, regulators need tools to identify potential benefits (e.g., "Physical Health" and "Liberty Protection") and harms (e.g., "Physical Injury" and "Liberty Loss"), and the magnitude of those impacts (e.g., likelihood and scope of impact) an AI system's development or use poses to society.[50] Accordingly, the White House Office of Science and Technology Policy (OSTP), in coordination with the Office of Management and Budget (OMB), or another equivalent government entity, should provide sector regulators with tools to determine which AI uses should be the focus of their regulatory efforts. This guidance should allow regulators flexibility to apply their sector-specific experience and expertise, but also be standardizable across agencies to provide industry and the public a level of certainty as to what AI uses will be considered highly consequential. Congress also should consider establishing a centralized AI authority that can regulate AI issues that cut across sectors and fill regulatory gaps in sectors.[51]

Regulators should expend their oversight efforts on AI use cases that are highly consequential to society.

48 See, e.g., The Dataset Nutrition Label, Data Nutrition Project (2023) ("The Data Nutrition Project takes inspiration from nutritional labels on food, aiming to build labels that highlight the key ingredients in a dataset such as meta-data and populations, as well as unique or anomalous features regarding distributions, missing data, and comparisons to other 'ground truth' datasets."). Model Cards report information about a machine learning model which can include its intended use, performance metrics, and limitations of the model. See Margaret Mitchell, et al., Model Cards for Model Reporting, arXiv (2019).

49 David Vergun, Defense Innovation Board Recommends AI Ethical Guidelines, U.S. Department of Defense (2019); Principles of Intelligence Transparency for the Intelligence Community, Office of the Director of National Intelligence (2015); OECD AI Principles Overview, Organisation for Economic Co-operation and Development (2019). On actions to operationalize responsible and ethical AI principles, see Key Considerations for Responsible Development and Fielding of Artificial Intelligence, National Security Commission on Artificial Intelligence (2021).

50 Such tools should be applied at different points in the AI lifecycle: (1) regulators foresee a new application for AI; (2) a new application for AI is under development or proposed to a regulatory body, and (3) an existing system has created a highly consequential impact that triggers a post facto regulatory review.

51 For an overview of governance options, see AI Governance Authority Options Memo, Special Competitive Studies Project (2023).

Digital Platforms from Countries of Concern. As we near the election cycle, there is an increased number of U.S. voters on foreign digital platforms from countries of concern. Concurrently, these platforms are converging with GenAI. For example, ByteDance has incorporated a chatbot "Tako"[52] into TikTok in Southeast Asia and is building out a large language model (LLM) for future use in its platforms under the codename "Grace."[53] GenAI presents the possibility of increasing the volume and speed of malign content on platforms. Moreover, GenAI increases the potential risks of sensitive data being used to target voters ahead of elections. GenAI models are typically trained on large amounts of data and given increased user engagement, privileged or sensitive user data is more accessible. The novel risk of these possibilities for digital platforms from countries of concern is that foreign governments or actors may have the ability to control or otherwise influence content, especially in cases where platforms are headquartered in locations where regulations and verification tools on platform use leave much to chance.

The United States should address threats posed by foreign digital platforms from countries of concern ahead of the 2024 election cycle using a two-path approach.

The United States should address threats posed by foreign digital platforms from countries of concern ahead of the 2024 election cycle using a two-path approach.

1. First, Congress should take necessary steps to consider narrow, product-specific restrictions on foreign digital platforms representing national security risks, such as TikTok.[54] A focused restriction would need to be introduced this fall for proposed enforcement at the start of 2024 ahead of U.S. elections. The ANTI-SOCIAL CCP Act[55] is an example of a legislative initiative with a narrow approach.

2. Second, the United States should also develop a more comprehensive risk-based, policy framework to restrict foreign digital platforms from countries of concern. The framework should consider a suite of legislative, regulatory, and economic options available to mitigate harm from such platforms. The framework should pursue a

52 Josh Ye, TikTok Tests AI Chatbot 'Tako' in the Philippines, Reuters (2023).
53 Zheping Huang, ByteDance, TikTok's Chinese Parent Company, Is Testing an AI Chatbot, Time (2023).
54 See Meaghan Waff, TikTok Is the Tip of the Iceberg: National Security Implications of PRC-Based Platforms, Special Competitive Studies Project (2023).
55 See S.5245, ANTI-SOCIAL CCP Act (2022).

comprehensive range of options that policy leaders could employ on a case-by-case basis and be developed simultaneously as policymakers introduce focused bans on platforms. The RESTRICT Act[56] is an example of a legislative initiative considering a broad, risk-based approach.

Governing Transnational Generative AI Challenges. GenAI's transnational nature makes international mechanisms a necessary corollary to domestic governance steps. GenAI's risks cut across borders, affecting both states' sovereignty and many shared societal equities from social harms to legitimate law enforcement needs. To support the development of a common international foundation around global GenAI implications, the United States should liaise with the United Kingdom to make a central output of the upcoming UK global AI safety summit[57] the establishment of a new multilateral and multi-stakeholder "Forum on AI Risk and Resilience" (FAIRR),[58] under the auspices of the G20. FAIRR would convene three verticals focused on:

GenAI's risks cut across borders, affecting both states' sovereignty and many shared societal equities from social harms to legitimate law enforcement needs.

1. Preventing non-state malign GenAI use for nefarious ends (e.g., criminal activities or acts of terrorism);

2. Mitigating the most consequential, injurious GenAI impacts on society (e.g., illegitimate discriminatory impacts due to system bias), and;

3. Managing GenAI use that infringes on other states' sovereignty (e.g., foreign malign influence operations or the use of AI tools in cyber surveillance).

FAIRR would convene relevant stakeholders — including national officials, regulators, relevant private sector companies, and academia/civil society — to work in a soft

56 See S.686, RESTRICT Act (2023).

57 UK to Host First Global Summit on Artificial Intelligence, Office of the Prime Minister of the United Kingdom (2023); Esther Webber, UK to Host Major AI Summit of 'Like-Minded' Countries, Politico (2023).

58 GenAI offers an initial focus upon which to form an institution like FAIRR. Over time, success could support expanding its remit to cover broader AI governance. For a fuller description of FAIRR's elements, see Appendix.

law[59] fashion toward interoperable standards and rules that domestic regulators can independently implement.[60] Operationally, a peer review process of domestic regulators would provide political pressure to abide by a commonly established set of terms.[61] Establishing FAIRR under the G20 – including the PRC[62] – would provide a sufficiently inclusive foundation to enhance legitimacy and provide global economic scope to drive wider compliance with the established rules. FAIRR should go beyond the core G20 states in determining its full voting members in two respects. First, FAIRR should include any nation-state home to a private GenAI actor's headquarters where the GenAI model sits above a certain compute threshold.[63] Second, FAIRR should include a voting representative from the qualifying private GenAI actors themselves in the multi-stakeholder vein of entities like the International Telecommunications Union.[64]

59 International conditions make a formal treaty unlikely, preferencing a soft law approach. See Anya Wahal, On International Treaties, the United States Refuses to Play Ball, Council on Foreign Relations (2022). On soft law, see Gary Marchant, "Soft Law" Governance of Artificial Intelligence, UCLA AI Pulse (2019); Kenneth W. Abbott & Duncan Snidal, Hard and Soft Law in International Governance, International Organization (2000).

60 Such a structure is similar to that of the Financial Stability Board (FSB). See About the FSB, Financial Stability Board (2020); Stavros Gadinis, The Financial Stability Board: The New Politics of International Financial Regulation, Texas International Law Journal at 163-64 (2013).

61 See Peer Reviews, Financial Stability Board (2021); Stavros Gadinis, The Financial Stability Board: The New Politics of International Financial Regulation, Texas International Law Journal at 160 (2013). The Financial Action Task Force (FATF) uses a similar peer review, or "mutual evaluation" approach. See Mutual Evaluations, FATF (last accessed 2023).

62 An approach that includes the PRC would be wise on both the merits of the issues and as a diplomatic consideration. See Annabelle Dickson, Lord of the Supercomputers!: Britain's AI Minister is a Hereditary Peer, Politico (2023) (quoting the new UK AI minister, Jonathan Berry, that supporting PRC engagement as "it would be absolutely crazy to sort of try and bifurcate AI safety regulation globally").

63 Compute threshold would be determined based on the actual number of operations applied during the model's training. The UAE's May 2023 unveiling of the 40 billion parameter Falcon 40B serves as a sample model that could yield inclusion for the host state. UAE's First LLM is Open Source and Tops Hugging Face Leaderboard, Wired (2023).

64 See Kristen Cordell, The International Telecommunication Union: The Most Important UN Agency You Have Never Heard Of, Center for Strategic and International Studies (2020).

Appendix:

The Forum on AI Risk and Resilience (FAIRR)

Background

While domestic measures are the required foundation for the governance of GenAI, its transnational nature makes international mechanisms a necessary corollary.[65] Domestic governance cannot function if actors can operate beyond a state's jurisdiction. GenAI tools cross borders with relative ease.[66] An international regime is required to close gaps. A series of tailored mechanisms are most appropriate as governance arrangements work best when fitted to specific issues and challenges. GenAI presents risks at the nexus of geopolitical, transnational, and societal concerns. This memo focuses on a subset of three interconnected aims:

1. Preventing non-state malign GenAI use for nefarious ends (e.g., criminal activities or acts of terrorism);

2. Mitigating the most consequential, injurious GenAI impacts on society (e.g., illegitimate discriminatory impacts due to system bias); and

3. Managing GenAI use that infringes on other states' sovereignty (e.g., foreign malign influence operations or the use of AI tools in cyber surveillance).

SCSP proposes a new multilateral and multi-stakeholder Forum on AI Risk and Resilience (FAIRR). That new regime must function within the realities of today's international landscape – deepening geopolitical tensions and limited trust. The world has met similar

65 While still in its early stages, there is a growing body of scholarship exploring international AI governance and historical precedents. See Michael Veale, et al., AI and Global Governance: Modalities, Rationales, Tensions, Annual Review of Law and Social Science (2023); Ian Stewart, Why the IAEA Model May Not be Best for Regulating Artificial Intelligence, Bulletin of the Atomic Scientists (2023); Peter Cihon, et al., Should Artificial Intelligence Governance be Centralised? Design Lessons from History, Proceedings of Proceedings of the 2020 AAAI/ACM Conference on AI, Ethics, and Society (2020).

66 For an example of concerns of how AI will proliferate across jurisdictions, see Brian Nussbaum, Offshore: The Coming Global Archipelago of Corrosive AI, Lawfare (2023).

challenges before. The Cold War superpowers recognized shared challenges and created tools to manage them.[67] More recently, in confronting transnational terrorism, climate change, and financial crisis, nations have sought to avoid ungoverned spaces as fonts of instability.[68] Therefore, FAIRR is the path forward as a new international entity scoped in mission and inclusive in nature.

Mission

FAIRR's work would center on three interconnected governance needs. First, nation-states share broadly common ground in preventing misuse and malign use of GenAI tools by non-state actors. Illicit uses of GenAI tools that contribute to harms from economy-sapping fraud[69] to bio-terrorism[70] are detrimental to all. Second, while the contours might differ state-by-state, countries generally agree on the need to prevent harms such as discrimination.[71] Governance alignment to prevent these harms from creeping across borders and negatively impacting populations would constitute a likely point of at least partial agreement. Third, states require a forum to discuss and deconflict instances where GenAI tools can cross boundaries and impact the sovereign activities of other states. For instance, the world will require a forum for dialogue to develop norms around legitimate GenAI-enhanced speech in an era of both growing misinformation and transnational censorship, potentially including around elections. Finally, pursuing all three sets of issues within one entity would help address the blurred lines between the three spaces. Particularly in the digital realm, states have deployed non-state actors for state ends.[72]

Structure

FAIRR would not constitute a new formal international agency, with international legal personality grounded in a treaty and creating binding rules.[73] Hurdles from treaty

67 U.S.-Soviet cooperation covered a range of issues from bilaterally collaborating on research in the health sciences to establishing the IAEA to curb the proliferation of nuclear weapons. See Bernard Gwertzman, U.S. and the Soviet to Pool Research in 3 Health Areas, New York Times (1972); Bertrand Goldschmidt, The Origins of the International Atomic Energy Agency, IAEA Bulletin at 15-16 (1977).

68 See Bruce Jones, et al., Power and Responsibility: Building International Order in an Era of Transnational Threats, Brookings Institution Press (2009).

69 Sam Sabin, Generative AI is Making Voice Scams Easier to Believe, Axios Codebook (2023).

70 See John T. O'Brien & Cassidy Nelson, Assessing the Risks Posed by the Convergence of Artificial Intelligence and Biotechnology, Health Security (2020).

71 While states can vary in their interpretation, wide acceptance of the International Convention on the Elimination of All Forms of Racial Discrimination speaks to a broad acceptance of this goal. See International Convention on the Elimination of All Forms of Racial Discrimination, 1969 (EIF), 660 UNTS 195.

72 Erica D. Lonergan, Cyber Proxies in the Ukraine Conflict: Implications for International Norms, Council on Foreign Relations (2022).

73 See Anne Burnett, International Organizations, American Society of International Law at 3-4 (2015).

ratification[74] to skepticism of international institutions[75] preclude that path. Instead, FAIRR would serve as a forum to promulgate alignment based on soft law.[76] FAIRR would convene relevant stakeholders — including national officials, regulators, relevant private sector companies, academia, and civil society — to work toward interoperable standards and rules that domestic regulators can independently implement.

Participation

Establishing FAIRR under the auspices of the G20 would both provide a legitimate foundation and establish a basis for its founding members. As with the Financial Stability Board (FSB),[77] FAIRR would benefit from being sufficiently inclusive to cover enough of the global economy to drive both member and non-member adherence to the established rules. Such an impetus would support seeking PRC participation in order to work toward globally comprehensive and consistent practices. Less ambitiously, even if the members became deadlocked, the PRC's participation in FAIRR would help avoid forum shopping[78] that could lead to the PRC rallying an international grouping under alternative rules inimical to U.S. values and interests.[79]

Furthermore, while drawing on a G20 foundation, FAIRR should voluntarily go beyond that core G20 membership in two respects. First, it should include as full original members any state with a headquarters of a GenAI actor over a certain compute threshold, such as the

74 For instance, the U.S. domestic political environment remains skeptical of ratifying treaties. See Anya Wahal, On International Treaties, the United States Refuses to Play Ball, Council on Foreign Relations (2022).

75 See Stewart Patrick, The Sovereignty Wars: Reconciling America with the World, Brookings Institution Press at 227-32 (2018); Julian Ku & John Yoo, Taming Globalization: International Law, the U.S. Constitution, and the New World Order, Oxford University Press at 42-47 (2012).

76 See Gary Marchant, "Soft Law" Governance of Artificial Intelligence, UCLA AI Pulse (2019); Kenneth W. Abbott & Duncan Snidal, Hard and Soft Law in International Governance, International Organization (2000).

77 See About the FSB, Financial Stability Board (2020); Stavros Gadinis, The Financial Stability Board: The New Politics of International Financial Regulation, Texas International Law Journal at 165 (2013). Robert Fay at the Centre for International Governance (CIGI) has led the way in exploring adapting the FSB structure to the broader digital domain. See Robert Fay, Global Governance of Data and Digital Technologies: A Framework for Peaceful Cooperation, Centre for International Governance Innovation (2022); Robert Fay, Digital Platforms Require a Global Governance Framework, Centre for International Governance Innovation (2019).

78 The concept of "forum shopping" in international scenarios builds on the concept of the domestic U.S. legal practice "of pursuing a claim subject to concurrent jurisdiction in the court that will treat the claim most favorably." Forum Shopping, Cornell Legal Information Institute (2022). On forum shopping in the context of international AI governance, see Peter Cihon, et al., Fragmentation and the Future: Investigating Architectures for International AI Governance, Global Policy at 550 (2020).

79 As the PRC's Regional Comprehensive Economic Partnership (RCEP) initiative illustrates, the PRC is more than capable of seizing open ground to propose its own standards and norms. See Michael Sutherland, Regional Comprehensive Economic Partnership (RCEP), Congressional Research Service (2020).

United Arab Emirates (UAE).[80] Such a move would help avoid nations with GenAI capabilities looking to caucus states to build alternative regimes. Second, as GenAI is first and foremost a private sector-driven advancement,[81] FAIRR should be multi-stakeholder in composition, including major private entities with GenAI capabilities over a certain compute threshold as full voting members. A weighted voting system[82] would help ensure national governments remained dominant over private entities in order to maintain the primacy of public interests. Additionally, an advisory committee(s) of academic and civil society experts could help balance private sector interests at the table.

Operations

Full members, original members (state and non-state), plus states that grow to meet the criteria of being home to an above-threshold model, would set relevant standards and rules. Members would participate in a double-weighted voting system to capture both state and non-state equities and disparate capabilities. Under this system, the total voting share of states would equal twice the voting shares of non-state members. Simultaneously, to distinguish among state capabilities, the total share of votes for states would be proportional to the number of qualifying private GenAI entities in their jurisdiction (with a minimum of one vote per state).

Similar to the FSB or the Financial Action Task Force (FATF),[83] a peer review process of domestic regulators would provide political pressure to abide by commonly set terms.[84] Noncompliant members would confront market ramifications based on their recognized higher risk, such as business withdrawals or chilled investment in response.[85]

---

80 In May 2023, the UAE unveiled a 40 billion parameter LLM, the Falcon 40B. UAE's First LLM is Open Source and Tops Hugging Face Leaderboard, Wired (2023).

81 See Harnessing the New Geometry of Innovation, Special Competitive Studies Project at 22-29 (2022).

82 While the FSB's Plenary operates on consensus, it offers a model in which countries receive different degrees of representation based on "the size of the national economy, financial market activity and national financial stability arrangements of the corresponding Member jurisdiction." Charter of the Financial Stability Board, Article 11 (2012). See also Diego Lombardi, The Governance of the Financial Stability Board, Brookings Institution at 10-11 (2011).

83 See What We Do, Financial Action Task Force (last accessed 2023).

84 See Peer Reviews, Financial Stability Board (2021); Mutual Evaluations, Financial Action Task Force (last accessed 2023). See also Stavros Gadinis, The Financial Stability Board: The New Politics of International Financial Regulation, Texas International Law Journal at 160 (2013).

85 As a comparable example, see the impact of FATF gray listing on foreign investment in Pakistan. Purvaja Modak, FATF's Scrutiny and What Non-Compliance Means, Centre for Public Policy Research (2021).

Funding

Resourcing must take place in a manner that maintains FAIRR's legitimacy. Perceptions of private funding leading to corporate interest capture would undercut the entire initiative.[86] Consequently, funding for the secretariat and operating costs should fall primarily (majority) with national contributions, likewise differentiated according to vote share. Private companies would contribute a minority share of the budget with proportions divided based on market capitalization.

86 Regulatory Capture, Oxford Reference (last accessed 2023).