FOUNDATION FOR AMERICAN INNOVATION

Dec 6, 2023

Written Testimony of Samuel Hammond
Senior Economist, Foundation for the American Innovation (FAI)
Before the
U.S. House Oversight Subcommittee on Cybersecurity, Information Technology, and Government Innovation, December 6, 2023

"White House Policy on AI"

Chairwoman Mace, Ranking Member Connolly and members of the Subcommittee, I thank you for the opportunity to testify today.

My name is Samuel Hammond, senior economist for the Foundation for American Innovation. FAI is a group of technologists and policy experts focused on developing technology, talent and ideas to support a freer and more abundant future.[1]

My research at FAI focuses on the second order effects of technologies like Artificial Intelligence on our institutions.[2] By second order, I mean not only what an AI system can do on its own, but what is likely to result as AI capabilities diffuse throughout the economy. These second order effects are all important, as the history of transformative technologies – from the printing press to the industrial revolution to the internet – is also a history of equally transformative changes to government.

I call the tendency to neglect the second order effects from technology the Horseless Carriage Fallacy, as if the advent of the automobile merely replaced horse drawn carriages while holding everything else constant. In reality, the automobile changed virtually everything, radically reshaping American institutions and economic geography.

Artificial Intelligence will do the same. The question is whether governments will keep up and adapt, or be stuck riding horses while society whizzes by in a race car. The risks from adopting AI in government must therefore be balanced against the greater risks associated with not adopting AI proactively enough.

---

[1] See: www.thefai.org
[2] "AI and Leviathan, Part I." Samuel Hammond. Second Best. (2023, August 23). https://www.secondbest.ca/p/ai-and-leviathan-part-i

It's within this context that I approached the White House's recent Executive Order on AI ("Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence"). A sprawling document, there is no doubt much to applaud in the Executive Order, including:

- Streamlining of visa processes to attract and retain immigrants with AI expertise,
- New initiatives to assess and address federal cybersecurity vulnerabilities,
- A "talent surge" to bring AI professionals into government,
- Exploration of Privacy Enhancing Technologies (PETs) to protect citizens' data,
- Discouragements against agencies instituting blanket bans on common AI tools,
- Chief AI Officers responsible for promoting AI in each agency, and an interagency council to coordinate the use and development of AI across agencies.

Dual-use foundation models

I was particularly impressed by the EO's focus on "dual-use foundation models," such as those that could be used to generate biological agents, cyberweapons, or similarly catastrophic hazards to the American public.[3] While "dual-use" can be a vague term, the definition used is narrowly targeted on relatively large models with specific capabilities, such as "enabling powerful offensive cyber operations," or "permitting the evasion of human control or oversight through means of deception or obfuscation."

The EO further requires the adoption of red-teaming and safety evaluations for the largest, most powerful AI models. It does this by using the authorities provided in the Defense Production Act, which may be controversial, but warranted insofar as the types of AI models in question represent bona fide risks to national security.[4]

This includes AI models trained on biological sequence data using 10^23 FLOPs, or any AI model produced with more than 10^26 FLOPs – a measure of the computing resources used in the AI's training stage. This threshold represents roughly two orders of magnitude more compute than was used to train GPT-4 – the most powerful AI model deployed to

[3] "Developing a National AI Strategy." Samuel Hammond. Comment in response to OSTP request for information. OSTP-TECH-2023-0007. (2023, July 7). https://www.thefai.org/posts/developing-a-national-ai-strategy

[4] See my discussion of AGI risk on the Future of Life Institute Podcast: "Samuel Hammond on AGI and Institutional Disruption." (2023, October 20). https://www.youtube.com/watch?v=AxrWNR3sBN0

date. No company has yet deployed a model large enough to meet this threshold, and likely won't for a year or more.

The primary shortcoming of a compute threshold is that dangerous AI capabilities do not necessarily correlate with the scale of the compute used in training. For that reason, the EO further establishes that the Secretary of Commerce will update as needed the set of technical conditions for models and computing clusters that would be subject to the reporting requirements. Nonetheless, compute remains a reliable proxy for the performance of generalist AI models,[5] and as such, the threshold is useful for picking out for special oversight the small number of companies attempting to create Artificial General Intelligence or AGI,[6] while leaving the vast majority of AI research and development unscathed. Note this light-touch approach is in some ways the inverse of the EU's AI Act, which imposes blanket safety and registration regulations across Europe's entire AI sector.[7]

As the computing resources used to train frontier AI models scale, they develop new and often unpredictable capabilities, including superior reasoning and planning abilities. Depending on how it's defined, the median Metaculus forecast for the arrival of AGI – a system that performs at human level or better at any task you throw its way – now ranges from 2026[8] to 2031.[9] Yet progress will not stop there. With the ability to control a robot or a computer's operating system and reliably perform complex sequences of tasks, including basic R&D, such generally intelligent AIs could unlock accelerating economic and scientific change – or wreak havoc if deployed by bad actors and without safeguards. Focused attention on AGI-specific risks is thus fully justified.[10]

---

[5] "We conclude that compute-based extrapolations are a promising way to forecast AI capabilities." Owen, David. "Extrapolating performance in language modeling benchmarks." Published online at epochai.org. (2023, June 9). 'https://epochai.org/blog/extrapolating-performance-in-language-modelling-benchmarks

[6] "Why AGI is closer than you think." Samuel Hammond, Second Best. (2023, September 22). https://www.secondbest.ca/p/why-agi-is-closer-than-you-think

[7] "America Cannot Afford to Be like Europe in Regulating Artificial Intelligence." Luke Hogg, The National Interest. (2023, June 22). https://nationalinterest.org/blog/techland/america-cannot-afford-be-europe-regulating-artificial-intelligence-206574

[8] "After a weak AGI is created, how many months will it be before the first superintelligent oracle?" Metaculus. https://www.metaculus.com/questions/4123/time-between-weak-agi-and-oracle-asi/

[9] "When will the first general AI system be devised, tested, and publicly announced?" Metaculus. https://www.metaculus.com/questions/5121/date-of-artificial-general-intelligence/

[10] "Frontier AI Regulation: Managing Emerging Risks to Public Safety." Anderljung, M. et al. Arxiv. (2023, July 6). https://arxiv.org/abs/2307.03718

Parallel requirements are extended to Infrastructure as a Service (IaaS) providers to report the identity of any foreign person who seeks to train a cyberattack-enabling model on their service. This includes resellers of their services as well. To reiterate, this is specific to training runs that exceed the 10^26 compute threshold, or meet the definition of "dual-use foundational model" above. Nevertheless, as an embryonic form of "KYC for compute," Congress should exercise its oversight powers to ensure such reporting requirements remain consistent with the law, do not become cumbersome, and resist mission creep.

Promoting AI in government

More than anything, the bulk of EO is aimed at promoting the use of "safe and trustworthy" AI within government while mitigating potential risks. It does this in a variety of ways, including through the formation of an interagency council to coordinate the use and development of AI in concert with the Chief AI Officers at each agency. It is difficult to assess in the abstract whether this broader framework will accelerate AI adoption in government or whether the new layers of oversight will simply slow things down.

Consider the EO's establishment of Artificial Intelligence Governance Boards with oversight over AI issues within each agency. On the one hand, this could expedite the adoption of AI by enabling a degree of central coordination. On the other hand, AI is increasingly embedded in every form of software, and often hard to distinguish from ordinary information technology. Which hand prevails will depend on whether the Boards and AI Officers spend their time accelerating the adoption of advanced AI or on policing the phantom risks posed by basic machine learning. As it stands, OMB was already two years late and well past its statutory deadline to issue AI guidance to agencies as required by the AI in Government Act of 2020.

The dangers from entrenching slow or duplicative approval processes is hard to overstate. FedRAMP was created to overcome this problem in the procurement of cloud services. As a government-wide compliance program, agencies can transact with FedRAMP-approved providers knowing that they meet a standardized level of security. The most common AI services should follow the same model and be evaluated and authorized for use government-wide, letting Chief AI Officers focus on overseeing bespoke or ad hoc AI systems for specific agency needs.

Winning the arms race

The case for aggressive adoption of AI in government comes down to the arms race between AI and our institutions. This is most obvious in the arena of cybersecurity. As AI lets hackers and other bad actors level-up their capabilities, our cyber defenses will need to level-up at least as fast. And yet these dynamics extend far beyond cases of explicit AI misuse. Democratized access to AI lawyers could quickly overwhelm the court system, for instance, just as expert AI tax accountants could soon democratize the ability for individuals and businesses to minimize or complexify their tax liability.

These vectors of attack don't constitute misuses of AI at all, but rather appropriate use at an unprecedented scale. In the near future, for instance, AI agents will likely fully automate the process of filing and appealing a FOIA request. Any information that can be requested will be, necessitating the adoption of e-discovery systems to allow AI agents to automatically review and fulfill the request on the government's end. The final equilibrium could make for a far more transparent and efficient government, but in the interim, having many more requests for information than an agency has capacity to fulfill could cause a de facto Denial of Service attack. Similar such "attacks" are likely possible across any number of public venues or services, from AI generated regulatory comments, to the sheer volume of economic activity unlocked by AI agents becoming illegible.

Even the most productive uses of AI will increase the throughput demand on our administrative state by orders of magnitude. Just last week, Google DeepMind published an AI model that discovered 2.2 million new crystals and 380,000 new stable materials that could power future technologies.[11] This represents nearly 800 years' worth of new material science knowledge, achieved virtually overnight. Imagine what will happen when this same pace of discovery comes to medicine, as it almost surely will. In a typical year, the FDA approves around 50 new molecular entities for novel drugs. Could the FDA handle increasing this approval rate to 500, 5,000 or even 50,000 new molecules per year, unlocking centuries of progress in personalized medicine? The answer is clearly no, at least not under business as usual.

[11] "Millions of new materials discovered with deep learning." Amil Merchant and Ekin Dogus Cubuk, Deepmind. (2023, November 29). https://deepmind.google/discover/blog/millions-of-new-materials-discovered-with-deep-learning/

In every case, managing these growing throughput demands will require the federal government to not only adopt AI aggressively, but should force Congress and the executive branch to rethink the configuration of our administrative and regulatory agencies from the ground up. From broken procurement policies to the bureaucratic sclerosis engendered by slow and outdated administrative procedures, incremental reform is unlikely to suffice. We must modernize government at the firmware-level, or risk ubiquitous system failure and government becoming *the* primary bottleneck to AI's enormous potential upside.[12]

Earlier this year, researchers at OpenAI published a paper assessing the likely labor market impact of Large Language Models. They found 80% of the U.S. workforce could have at least 10% of their work tasks affected by the introduction of LLMs, with jobs like Accountants, Auditors, and Legal Secretaries facing an exposure rate of 100%.[13] Many large companies have already begun downsizing or have plans to downsize, in anticipation of the enormous efficiency gains unlocked by emerging AI tools and agents.

Much of the work performed in government bureaucracies is especially low-hanging fruit for AI.[14] With just under 3 million employees in the federal workforce, Congress should demand the White House and OMB undertake an analogous survey to discover which federal jobs are most exposed to AI, and to what extent legislation is needed to expedite new, AI-enabled models of governance. The goal should not be to downsize the federal bureaucracy per se, but rather to augment employee productivity and free up human resources for higher value uses, reducing waste and enhancing capacity simultaneously.

Take the FTC's health care division, which employs around 30 attorneys to police competition across the entire U.S. health care industry. A day in the life of these attorneys looks like manually reading through tens of thousands emails subpoenaed from a pharma CEO as part of discovery. Yet today, with the right prompt engineering, one could feed those emails into a Large Language Model and simply ask it to find the most egregious examples of misconduct. This wouldn't replace the attorney's role in verifying what the AI discovers, but even with the imperfections of current models, it would nonetheless drive massive productivity gains – gains that we can be sure are being exploited by the private law firms on the other side.

[12] "Heretical thoughts on AI" Eli Dourado. (2023, January 19). https://www.elidourado.com/p/heretical-thoughts-on-ai
[13] GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models." Eloundou, Tyna et al. OpenAI. (2023, March 17) https://arxiv.org/abs/2303.10130
[14] "Disrupting Bureaucracy" Samuel Hammond. PlainText. (2015, September 24). https://readplaintext.com/disrupting-bureaucracy-fa611d04f956

Given bureaucratic inertia, it is not enough to simply ask agencies to prioritize the use of generative AI in government, as prescribed by the Executive Order. Congress must push the federal government to move faster, including by authorizing additional AI training resources[15] and funding for technological modernization.

At the same time, the same tools that can be used to enhance federal capacity can be further used to strengthen Congressional oversight. Most of the work and communications performed in any given agency is now machine readable. As agencies embrace AI internally, managers will be able to easily track and query the performance of their staff, automatically generating reports and work summaries from common document repositories. These same techniques could be used to expedite reports to Congress, and even enable near real-time monitoring of an agency's activities. Call it Inspector General-GPT.

Innovating within government should mean more than plugging AI into an existing, outdated process and calling it a day. It will take true inventiveness and ambition. So while the White House has made some important first steps, there is much more to be done. With appropriate urgency and coordination between the Executive branch and Congress, we can forestall system failure by co-evolving our institutions with AI, enhancing public trust and saving taxpayers' dollars in the process.

Thank you and I look forward to your questions.

Samuel Hammond
samuel@thefai.org

[15] See, for example, the *AI Training Expansion Act of 2023*. https://www.congress.gov/bill/118th-congress/house-bill/4503