TRANSCRIPT
AI Policy Podcast
“AI 2024 Outlook”

DATE
January 23, 2024

CSIS EXPERTS
Gregory C. Allen
Director, Wadhwani Center for AI and Advanced Technologies, CSIS

H. Andrew Schwartz
Chief Communications Officer, CSIS

Transcript By
Rev Transcript
www.rev.com
Gregory C. Allen: Welcome to the AI Policy Podcast, a podcast by the Wadhwani Center for AI and Advanced Technologies at CSIS. I'm Gregory C. Allen.

Andrew Schwartz: And I'm Andrew Schwartz.

Gregory C. Allen: Join us as we dive into the world of AI policy where we'll discuss the implications of this transformative technology for national security, geopolitics and global governance.

Andrew Schwartz: And we're back on our brand new AI policy podcast and we want to talk this week about AI in 2024. Greg, as you think about AI in 2024, what developments in AI tech itself do you expect to see?

Gregory C. Allen: I think one thing that we've actually been riding in the current wave of AI technology is what is called the scaling hypothesis. And this is an empirical observation about how AI systems improve. It turns out that if you just throw a lot more data and a lot more computing resources at existing approaches to training AI models, the models that come out with more data, with more compute, just keep getting better. And that has been true for a long, long time at this point. And in 2024, we're already sort of running up against what might be the limits of what the capital markets can actually provide in terms of the money to train these big AI models. So for example, GPT-4, which is the most advanced AI system, large language model that's available to the public right now, that system OpenAI has disclosed costs something like $150 million just to build the data centers and to use all the electricity to run all the computations to analyze all the data.

Well, the next system after this might cost something like a billion or a billion and a half just to train this AI model. And then it costs a lot more when you want to use it. Every time you're entering a search into Google, it costs money for Google to run the computations to go find the things that you want in Google search, but it's like fractions of a penny. When you're using AI for large language models, it costs a lot more than that. So the costs of building and operating these systems just keep going up.

And so I think there's two questions that we have for 2024 - how many people are actually going to be able to put up the kind of money that they need to build these things when we're talking billions of dollars for the most advanced AI models? And then second, is the scaling hypothesis going to stay true? If you increase the amount of computing power, if you increase the amount of data by 10 x, does the thing get 10 times smarter or how much smarter does it really get? And is the juice worth the squeeze in expending all those additional resources?
Andrew Schwartz: So given all this, do you see major breakthroughs in 2024?

Gregory C. Allen: I think definitely we’re going to see performance go way up. I do think it’s sort of debatable whether or not 24 is likely to bring a breakthrough. From my perspective, the real breakthroughs that we’ve seen in AI, and I’m talking here, really transformative breakthroughs, was sort of in 2012 when we figured out how to connect deep learning neural networks to GPUs. We were sort of riding that wave for the next eight to 10 years, just applying that approach to machine learning with that computing hardware. And then in the past few years, you have this new revolution centered around large language models and what is called the transformer architecture, which is sort of a specific way of designing machine learning AI systems. And I think that architecture still has a few more years of juice to squeeze out to keep with that metaphor. And then the question is, when is there going to be revolutionary new architecture on top of that?

If you listen to folks like Sam Altman, the CEO of OpenAI, they say that if you want something like artificial general intelligence and AI that is as smart as a human being in most of the ways that human beings are smart and capable from an economic perspective, that is going to require a new architecture. And so, transformers are awesome. Large language models are awesome. There’s plenty of ways to deploy that in the economy and society that are still going to have, but you also have some of the smartest people in the world working on the next revolution, and it’s really tough to say when that’s going to come. Could come in 2024, could come in 2029,

Andrew Schwartz: But is there any question in your mind that the investment is going to go into this?

Gregory C. Allen: Oh, absolutely. I mean, what I’ve heard from the venture capital community is that interest rates are really high. It’s really hard to raise money for a new startup unless you’re working on AI. There’s money for AI.

Andrew Schwartz: That changes everything.

Gregory C. Allen: Yes, exactly. And that’s true in startups and that’s true in big enterprise companies. Everybody is pouring tens of billions of dollars investment into AI, and that’s definitely going to lead to some breakthroughs of some kind.
Andrew Schwartz: Okay, so before we dive into your AI policy predictions for 2024, I want to touch quickly on elections because 2024 is going to be the biggest election year in our history. What role do you expect to see AI play? And we've already seen AI generated misinformation in elections abroad. Do you expect that this trend is going to continue in U.S. presidential and congressional election?

Gregory C. Allen: Yeah, so 2024 has almost every democracy on earth, just coincidentally has an election in 2024. The stars have aligned, the calendars have aligned, and so there’s so many elections in 2024 and now AI is going to be a factor in two ways. Number one, I think AI is going to be part of the machinery of campaigning and running elections. We’ve already seen some of this in the United States. Ron DeSantis released a chatbot where you can chat with Ron DeSantis and of course it’s a large language model AI under the hood. I don't know that it's actually persuading many voters, but this is capabilities that campaigns are exploring. Maybe somebody will figure out how to use it in an impactful and serious way.

The other part of this story is around deep fakes and misinformation, because AIs can create photos that look extremely realistic. They can create audio and video that look extremely realistic, and with social media platforms, it’s possible for that media to circulate in front of a lot of folks before the fact-checkers sort of get their pants on.

This is something that we’re already reportedly seeing in the Bangladeshi election in 2023, where reportedly Deepfakes are a big part of the story, leaking false audio of your political opponents saying something incredibly inflammatory or horrific and doing the same with images. So, this is already happening in parts of the world, and it might come to the United States, it might come to other advanced democracies. The question is, are there going to be effective safeguards on that? I assume that the Federal Elections Commission will come up with effective regulations that deal with the specific campaigns for the Biden campaign and the Trump campaign, but Super PACs, other sort of secretive organizations, all of these could do a lot with this AI technology and perhaps not be as effectively regulated as the main campaigns themselves.

Andrew Schwartz: So we’re talking about campaign ads or campaign messages of some sort?

Gregory C. Allen: Well, both the information that they circulate, which might be disguised as news, it might be released into the parts of the news media and social media ecosystem that don’t have a lot of fact checks. I mean, we saw some of this in the 2016 election where the false
article that the Pope had endorsed Donald Trump got tens or hundreds of millions of views even though it was completely fake. I think the question then becomes like, well, that was just an article. Right? What if it is a high definition video that looks quite compelling of Joe Biden insulting whatever political demographic you want to say? A lot of people might see that video on social media or elsewhere, and they might not later hear or they might not believe the Biden campaign’s denial and this sort of thing is an experiment that we’re probably going to run in 2024 whether we like it or not.

Andrew Schwartz: Wow, that’s a lot to think about, and I’m sure this is something we’re going to be talking about in an ongoing way on our podcast. And by the way, we’re going to bring in other guests to this podcast too, some of the experts who cover this in the media, some of the experts who we can learn from and have a good discussion with. So, we’ll look forward to that in 2024, but let’s talk about your policy predictions for 2024. Let’s first start with regulation. 2023 saw several high-level AI governance efforts unfold. Do you expect these efforts to translate into actionable policy in 2024?

Gregory C. Allen: I think we should start with the White House Executive order because, like a lot of executive orders, it was the White House telling the federal agencies to do something. So, in 80 days, you must come up with standards for AI uses in blobbity blah, or in six months you must develop regulations for blobbity blah. So, a lot of that sort of outsourcing from the White House to the federal agencies, that stuff’s going to start coming back and the Department of Commerce in particular is going to have a really big role to play here. They’re the ones who’s going to be overseeing these red teams for the foundation frontier models for the most advanced AI systems. But most agencies are going to have to come up with regulations for how AI is specifically affected in their domains, like the Department of Transportation, the Department of Energy. Really nobody is spared from having to do that, and that’s definitely going to come out in 2024.

Andrew Schwartz: What about Congress? Will they pass comprehensive legislation?

Gregory C. Allen: So comprehensive legislation is definitely the goal. Senate majority leader, Chuck Schumer has repeatedly said that. The problem is what we were just talking about. 2024 is an election year and it’s really hard to pass comprehensive legislation in an election year. It’s really hard to just pass the federal budget in an election year. So I think one thing that might happen here, and I’m totally speculating, is I think they are going to come up with a bill. I think they are going to come up with a lot of draft legislation and there’s going to be a lot of
negotiation about what that draft legislation looks like. And then when might it pass? If I had to guess, I would say the lame duck Congress. So right after the election, the political stakes are a little bit lower. You’ve got some people who are going to be out of office, they’re going to retire. The pressure on them to vote a certain way might go down, and so I think the most of the year is going to be negotiating what’s in that comprehensive package. And then the lame duck Congress would be my best guess for when something might actually pass, if something is indeed going to pass in 2024.

There’s one possible alternative there, which is maybe AI is a big factor in the election as an issue. I mean, most people who vote are voting on their economic livelihood, their cultural identity, whether or not they feel safe. The question is, is there really a meaningful group of people out there who would care what candidates’ positions are on AI policy? If I had to guess, I would say this is going to come up in presidential debates. It already has in the presidential primary debates. But -

Andrew Schwartz: If we actually have presidential debates.

Gregory C. Allen: Yes, that’s a totally fair point. But if we have presidential debates, it’s totally plausible to me that you can imagine one of the questions being asked related to what are you going to do about AI? And so, could AI be a significant issue in the elections? Could that actually affect what does or does not get passed? I think that’s plausible.

Andrew Schwartz: And so, what can we expect to see globally? There’s going to be a G summit in June in Italy. What do you think they’re going to do in AI?

Gregory C. Allen: So, Italy is now the next rotating president of the G7. Japan was the G& president in 2023. Italy is the president in 2024, and I think Italy is now coming into this with the EU AI Act behind it. So, Italy in the past has sort of been at the more extreme end of restricting AI. When Chat GPT first launched, it was actually banned in Italy briefly, but it was the Italian business community that came in and said, ‘Hey, this is the next technological frontier. We need to be able to use this. We don’t want to hamstring our own economy.’

And so, what’s interesting is at the same time the EU AI Act is now moving into the implementation stage, a European country, Italy is now going to run the G7, and so I think there’s that opportunity for the conversation for Italy to seize. I’ve also heard that AI among Italy’s priorities for the G7 might not be their highest priority right now. The current regime in Italy is really interested in topics like immigration. But because there is this Digital and Technology
Ministers’ Summit that always occurs within the G7, obviously the folks who do AI policy are going to meet, presumably they’re going to talk about AI policy.

Andrew Schwartz: So, if you had one AI policy development that you hope to see take place in 2024, what would it be?

Gregory C. Allen: Great question. I think the thing that I would really want to see is around actually AI and national security. So we’ve talked so much about AI regulation and governance because that is the hottest topic, but when I think about what is shockingly not working yet - it’s the U.S. military’s adoption of AI and the replicator initiative that we talked about in the last podcast is still unfunded. I would love to see that actually get money behind it.

Andrew Schwartz: Let’s talk about AI and national security. How do you think the DOD’s big AI and autonomy bets, like Replicator and Task Force Lima will play out next year?

Gregory C. Allen: So, I think the Department of Defense is reacting to two things. The first is what we’ve seen in conflicts, like Ukraine where they know that AI has a lot of military potential and they want to harness that, and I think replicator sort of reflects that insight. Then you have taskforce Lima, which is much more about generative AI - stuff like Chat GPT, stuff like stable diffusion image generation. The taskforce Lima is sort of exploring and figuring out what’s the right thing for the Department of Defense to focus on those two areas. In terms of replicator, I think the big question is funding and then what are they actually going to buy?

So, in the case of replicator, Deputy Secretary of Defense, Kathleen Hicks has sort of set this 18 to 24 month deadline. And typically, that means you’re going to buy stuff that already exists, right? When you say stuff is going to be in the hands of war fighters in 24 months, usually in the Department of Defense, that means you’re going to buy something where it’s already been built, it’s already been tested, and all we need to do is make a lot of those things.

So, what is the DOD going to buy and who is actually going to be responsible for it? So far, deputy Secretary Hicks has indicated that the Defense Innovation Unit led by former Apple executive Doug Beck are sort of in the driver’s seat, but historically, DIU hasn’t had a lot of money either. And so there’s this big question about how are you actually going to light this fire of change under the Department of Defense if the direction that you’re telling everybody to move in doesn’t have any money behind it.
Let's talk about AI enabled weapon systems and how will they continue to play a significant role in the conflict in Ukraine? This is real world stuff.

Yeah. The war in Ukraine has moved off of the headlines a little bit as the war in Gaza has really moved to the front, but it's still going on. It's still the largest land war in Europe in many, many decades, and it's a war where AI technology is already playing a big role. So the Ukrainian military has gotten a lot of attention for their use of satellite communications from stuff like SpaceX Starlink. But they're also pretty good at just doing software in general. I've heard from NATO leadership about their admiration for Ukraine's,Uber for artillery, where you can sort of just say, 'that's a target I want artillery on.' And then the matching algorithm will identify all the artillery that the Ukrainian military has within range, decide who is best positioned, and then issue the order to go actually initiate that strike.

Thus the Uber.

Thus the Uber, right, for artillery. That's sort of a simple rules-based algorithm. But Ukraine has also demonstrated a lot of skill in doing image recognition and computer vision machine learning AI for stuff like target recognition. And that so far has mostly been confined to intelligence surveillance and reconnaissance applications.

I think 2024 is probably the year where we're going to see widespread use of lethal autonomous weapons. If I had to guess. In the case of Russia, they have weapons manufacturers that advertise these capabilities that it can go fully autonomous. It does have AI computer vision on board, and you can use it for offensive strikes. This is like the marketing brochure for these weapons systems and those weapons systems are being deployed to Ukraine. And then on the Ukrainian side, the Minister of Digital Transformation stated explicitly that they thought that autonomous weapons were a logical and inevitable next step. So, when your country's existence is on the line, as is the case for Ukraine, certain types of weapons seem like they're worth using. This is fundamentally going to become a gloves-off conflict if Ukraine feels like its national existence is threatened. And so my guess would be 2024 is when we're going to see lethal autonomous weapons systems used.

So there's real potential here for an escalation in that regard, isn't there?

If you think this is an escalation? Yes. Right. There have been reports of these systems being used in other wars like the war between
Armenia and Azerbaijan not that long ago. I’m actually a little bit skeptical that the reported instances were a true offensive use of AI enabled autonomous weapons. So this is kind of a threshold that humanity is about to cross if we have not already crossed it. It’s possible that it’s going on in Ukraine and Russia and they’ve just managed to keep it under wraps for whatever reason. But if we haven’t already crossed that line in 2023, I think it’s pretty likely that we’ll cross it in 2024.

Andrew Schwartz: Alright, here’s another touchy subject. Congress still hasn’t passed the fiscal year 2024 budget. How might this affect AI and autonomy development in DOD?

Gregory C. Allen: Okay, I want to answer your question with a question. Andrew, I’ve got this really exciting new policy. Okay. It is a law I want to pass. The law just says that we’re going to make every person in the DOD and every organization in the DOD 10% dumber. Should we do it?

Andrew Schwartz: Yeah, probably not.

Gregory C. Allen: Probably not. And yet we do that every year when we pass a continuing resolution. It literally just makes every single thing you do harder because for those who are not familiar with a continuing resolution, that’s where the government says, ‘Hey, we can't agree on what the next year budget is going to be. So for the time being just keep doing the last year's budget.’

Andrew Schwartz: And we’ll kick the can down the road.

Gregory C. Allen: And we’ll kick the can down the road, which, in a company might not be the end of the world, but in the government, it is a violation of law to use your budget for something other than what it was stated to be used. So if you have already bought all the tanks that you need to buy and then a continuing resolution happens, turns out you’re buying more tanks even though you don’t want ‘em, even though you don’t need ‘em.

This is what I mean when I say that a continued resolution is basically a policy of just making the DOD dumber and it is productivity kryptonite for the AI and autonomy community because most of what they want to do is new. It's a new technology. And so when you pass a continuing resolution, you're saying nothing new. And that means you're often saying, no new AI, no new autonomy.

Andrew Schwartz: So CR could be really bad for the future of AI development at DOD.
Gregory C. Allen: It’s always really bad. When I was in the DOD, it happened to us and it was just so infuriating. It’s really remarkable that Washington spends so much time trying to think about how they could improve national security with this policy, how they could improve national security with that policy. There’s probably no idea being debated in Washington this year that could help national security more than just reliably not having continuing resolutions. You can tell I care a lot about this.

Andrew Schwartz: Yeah, fascinating. Well, we’ll watch that for sure, and definitely something we’re going to be talking about. Let’s shift to AI export controls and our competition, U.S. competition with China. In 2023, the Biden administration worked hard to restrict China’s access to AI chips, and should we expect to see more restrictions. Even more restrictions in 2024?

Gregory C. Allen: So I think there’s two things that really are on the table for 2024. The first is just updating the regulations. So the regulations were just updated in October, 2023. So there’s not sort of a desperate need to update them at this moment. But what the Biden administration communicated pretty clearly is that they recognize that technology’s moving fast. When you draw the line at one technology performance threshold, one year, engineers will work around that in the next year. So they’ve actually stated that they intend to update this policy on a minimum yearly cadence. So we could see updated export controls in June, we might see them in October, who knows? But we’re definitely going to see some kind of update.

The second thing that really needs to happen is around making the export controls more multilateral. So right now, Japan and the Netherlands are sort of the two big dogs in semiconductor manufacturing equipment other than the United States, and that’s now restricted to China.

The problem is that while the Chinese semiconductor manufacturing equipment industry is way behind Japan, the Netherlands and the United States, other countries like South Korea are only a little bit behind. And so when Chinese companies show up to South Korea, I’m exaggerating here, but basically with trash bags full of money and say, ‘Please sell us all the equipment you can sell us and please teach us how to make the equipment that you can make.’ Suddenly, China might be 15 years behind the state-of-the-art in the Netherlands, but South Korea is maybe five years behind. And so that makes a big difference when they can get that kind of help.
So as you can imagine, the Biden administration is really looking to get the other big countries in this area. That’s places, especially South Korea, but also like Germany, which does not do much in the way of making the finished machines for making these chips, but does a lot for making the components that go inside them. And all of that is just accelerating China’s path to de-Americanization and to building the machines that can build the AI chips that can help them win the AI competition with the United States.

Andrew Schwartz: So there’s a lot of pieces to this puzzle, but what do you think success looks like for the Biden administration in 2024 with export controls?

Gregory C. Allen: So I think the opportunities that were perhaps available in 2018 when the Trump administration set the United States on this journey of semiconductor export controls against China - the opportunities that were there, perhaps in 2018, they’re not there anymore. In 2018, if the United States, Japan and the Netherlands got together and restricted advanced equipment, China’s not making advanced chips. Period.

Now in 2023, they’ve already bought a lot of this equipment. They bought it when the original export controls from the Trump administration had a lot of loopholes so that if it was shipped through a third-party country, it wasn’t illegal anymore. They’ve bought it because the Dutch and the Japanese regulations took a long time before they came online. And so you’re basically dealing with the problem of China’s equipment stockpiling. There’s already a ton of this stuff in China. So in the case of Huawei and Smick, which are the sort of team of two companies who are the furthest along in designing and manufacturing AI chips in China, they already have enough equipment in China to I think produce something like 35,000 wafers per month. A wafer being the big silicon disc that you print computer chips on, and then you cut into squares to get the computer chips out. So that’s a lot of chips. That’s not nearly as much as China wants.

To finally come back to your original question about what does success look like? Success probably does not look like China never makes a seven-nanometer advanced AI chip, but it might look something like they produce only 10% of what they wish they were producing. So we probably cannot stop it, but we can definitely make it super expensive, super complicated, and at the end of that story, you don’t have nearly what you want.

Andrew Schwartz: So do you expect to see any new policy instruments deployed and developed to prevent China’s access to AI chips?
Gregory C. Allen: We've already seen some new tools be added to the toolbox. The outbound investment screening that came in 2023 where the US government through a White House Executive Order gave itself permission to review and deny investments into China's advanced technology ecosystem, including AI and semiconductors. That was a new power that the White House gave itself, and the Treasury Department has already sort of figured out how they're going to use certain parts of this power. For example, I don't think anybody in the United States is ever going to be investing in Chinese semiconductor manufacturing equipment ever again. But on other parts of that, like Artificial Intelligence, the Treasury Department did not provide firm regulations and instead was asking questions about how could we effectively draw these lines to cut off all the military types of AI that we don't want U.S. investors and U.S. capital markets involved in, but perhaps allowing some more of the innocuous commercial AI. It's a really hard way to draw that line, and I don't think the government knows how to draw it yet. But in 2024, they'd better have an answer because the timelines for when they're obligated to come up with these regulations take place in 2024.

Andrew Schwartz: Alright, so what do you think happens to China's AI ecosystem in 2024?

Gregory C. Allen: So, China has been number two worldwide in global AI in terms of both research - they're generating really high quality research and they're generating a lot of it - and also commercial adoption. They have a lot of great technology giants who were actually using AI to make money. So China's been number two behind the United States for quite some time now. And these export controls were designed to stop that, not immediately, but eventually. The problem is that China has stockpiled a lot of these AI chips. So NVIDIA announced in the middle of 2023 that they had $5 billion worth of orders from Chinese technology companies for the AI chips that they were legally allowed to sell to China. The updated export control regulations that happened in October will interrupt some of those sales because the deliveries hadn't been completed yet, but it won't stop all the sales that had already happened.

So a company like Tencent, which is the Chinese maker of WeChat, their sort of dominant social media platform and a very skilled AI company in a lot of ways, they announced that they have enough AI chips they think to meet all their needs for the next two years. So the problem from the perspective of the Biden administration is they adopted these really tough export controls, but some of the companies that they were going after have already bought all the chips that they're going to need for 2024. There will be some Chinese companies who are probably left out to dry and will fall behind in the
most advanced AI research, but Tencent is definitely not one of them. There’s probably others who stockpiled.

Andrew Schwartz: Okay. So how do you assess Huawei in 2024? What are their prospects?

Gregory C. Allen: Huawei is, I wrote a report about this in October, and I do think it’s appropriate to say at this point that Huawei is leading team China. They’re in a very privileged position to set semiconductor policy in the Chinese government. They’re sort of the official leader of the industry group that the Chinese government is backing. And so while they’re a privately owned company, in a lot of ways they look like an extension of the government and certainly the government is backing them with everything they have.

Huawei has advanced chip design capability. They were one of the first companies in the world along with Apple to introduce seven nanometer chips in smartphones back in, I think it was 2019. And they designed those chips in-house. They didn’t outsource it or something like that. And it’s very clear that in the time that they’ve been subject to export controls and cut off from advanced chip manufacturing, the quality of that design team has not degraded at all. They’re still good at that.

So they’re best known for their smartphone chips, but they actually have a lot going on in AI chip designing, and this is their Ascend product line. And what they’re really hoping to do is to break the NVIDIA dominance of the AI chip market. So prior to the Biden administration export controls, NVIDIA was estimated to have like a 95% market share in China for the chips that are used to train AI models. And the Huawei products in this area have always been inferior on all the performance metrics that matter. But now with these export controls, you’re not allowed to buy the best NVIDIA chips. And so even though the Huawei chips are worse, the question is how much worse are they than the NVIDIA chips that Nvidia is actually allowed to sell? And how many of these chips can Huawei make? And that’s what they’re going to try and do in 2024 is they’re going to try and make the best chips they can make and they’re going to make as many of them as they can.

Andrew Schwartz: Greg, this is a fascinating discussion and one I’m really looking forward to having with you on an ongoing basis in 2024. Thanks so much for your insight. We’ll be back in a couple of weeks with more on the AI Policy Podcast.

Gregory C. Allen: Thank you.
Thanks for listening to this week’s episode of the AI Policy Podcast. Be sure to subscribe on your favorite podcast platform so you never miss an episode. And don’t forget to visit our website csis.org for show notes and our research reports. See you next time.

(END.)