THE BROOKINGS INSTITUTION

WEBINAR

The geopolitics of generative AI

Wednesday, July 19, 2023

PANEL DISCUSSION

MODERATOR: JESSICA BRANDT
Senior Fellow and Policy Director, Artificial Intelligence and Emergency Technology Initiative, Brookings

SAMM SACKS
Senior Fellow, Paul Tsai China Center, Yale Law School
Cyber Policy Fellow, New America

MARIETJE SCHAAKE
International Policy Director, Stanford University Cyber Policy Center
International Policy Fellow, Stanford Institute for Human-Centered AI

CHRIS MESEROLE
Senior Fellow and Director, Artificial Intelligence and Emergency Technology Initiative, Brookings

* * * * *

BRANDT: Hi, welcome everyone, and thanks so much for joining us for this afternoon's event on generative AI and geopolitics, how it will both shape geopolitics and be shaped by geopolitics. We have a lot of threads to pull and fortunately, some terrific experts here to help us to do that. Samm Sacks is a senior fellow at the Paul Tsai China Center at Yale Law School and a cyber policy fellow at New America. Her research examines China's information and communications technology policies focused on the U.S.-China technology relationship and the geopolitics of data privacy and cross-border data flows. So we'll be turning to her for insights on how China is doing this AI moment. Marietje Schaake is the international policy director at Stanford University's Cyber Policy Center and an international policy fellow at Stanford's Institute for Human Centered AI. She's a former member of the European Parliament where she focused on trade, foreign affairs and technology policy. So no doubt you'll help us to better understand the landscape from European perspective. And my colleague Chris Meserole here at Brookings is the director of our AI and emerging tech initiative and a fellow in the Foreign Policy program. And his work focuses on how democratic societies and in particular the United States, should respond to the exploitation of digital technologies by authoritarian regimes. So I think we're poised for a very rich discussion.

I'd like to get all of you online engage, too. So I hope you'll join the conversation by submitting questions to events at Brookings dot edu excuse me, or by Twitter using hashtag geopolitics AI. And I and I know a number of you have done that already, so thank you very much. For now, let me start off by turning to you, Samm. Can you give us a sense of what is the state of play when it comes to generative AI in China? You know, we know that when it comes to

foundation models, Chinese, Chinese achievements have been lackluster. And I'm curious if you can help us understand why and whether this trend is going to continue.

SACKS: Sure. Thanks so much for having me. I also think it's interesting that so much of the discussion about AI governance and development is happening through the lens of geopolitics. And I hope that we can also sort of back up at some point in our discussion and talk about where are these China and geopolitical questions and where are these broader questions about how this technology is shaping all societies in ways that have nothing to do with China? But I'm happy to dive in first on the China lens. There's been a lot of discussion comparing China's achievements so far with the U.S. I think that the sort of conventional view has been that China is around 2 to 3 years at least behind U.S. counterparts, with advances in Chinese AI relying on iterations of a cutting edge research that's published abroad. In a recent Foreign Affairs piece, and it was described as - by Helen Toner, Jenny Xiao, and Jeff Ding, which I highly recommend everyone reading - they described it as a sort of drafting behind a slower cyclist, drafting behind the leader. And there's been, you know, even just in the past week, New York Times, for example, has compared Ernie Bot, which is Baidu's sort of counterpart to ChatGPT. And so, you know, there is a lot of discussion around the pros and cons of where everyone is.

I see China facing a few headwinds. The first and I think the most important one that we're going to have to watch for the party leadership really grapples with is the question around censorship and how to balance information control with China's ambition for global leadership in AI. Because AI generates and disseminates information that's a real concern to the Communist Party leadership and how they navigate this is going to be really telling. Under a generative AI regulation, which takes effect August 15th, which we're going to talk more about in the discussion, companies are liable for the for the content that the generative AI produces. And so one of the questions that I've had is how the how is the leadership going to calibrate this quest for control in relation to the quest for innovation and leadership? Recently, there have been two developments that I think are significant. When the government first unveiled these regulations, they were very burdensome in terms of providers and what they would be liable for in terms of content In a revised version of the regulation that came out very recently, they appeared to loosen this restriction in a way by by limiting the scope of it just to public facing-AI. And so that would provide potentially more space with less sort of burdensome requirements around content.

The other thing that I'm watching is, right now, the Cyberspace Administration of China, which is China's cyber regulator, has really been in the driver's seat. They drafted this regulation. But we know that the Ministry of Science and Technology is likely going to have the pen for China's AI law, which is in the works. The CAC, if they had their way, I've been told, would lock it all down. They are a propaganda-driven organization where security and control are at the heart of their mission. But with the Ministry of Science and Technology stepping in around the AI law, that could be an effort to sort of offset some of the focus on control around information to fuel innovation. So that's been I'm watching. Related ahead, one has to do with data because it's censored online environment really does have a limiting effect on the availability and the quality of data. And I think already data is a constraint on generative AI models globally. The other - and this is very much in the geopolitical lane and then I'll wrap up quickly - is Compute. The U.S. has restricted China's access to the most advanced semiconductors, and Chinese LLMs we know rely on advanced chips that are developed by the U.S., with Chinese semiconductor industries generally several generations behind. And so at the moment there's a loophole where Chinese companies have been using cloud service providers to rent access to some of these most advanced chips, including the A100s, which were restricted by the U.S. So we need to watch. Is this an area where the U.S. government is going to come out and try to close the loophole? How successful will China's AI ambitions be given these constraints coming from the U.S.? And I'll stop there.

BRANDT: Thanks, that's really helpful. And I'm glad you mentioned the Compute issues. I think that's something we're going to come back to; I'd love to draw Chris out. But before we move on, I just I would love to draw you out a little bit more on on the question of China's AI regulations. You know, you mentioned a bit, but I would love to just a little bit more detail on what do you think

is most sensible. What's sensible, I guess, in in China's approach and where do you see that, you know, it's sort of government's tradition of heavy-handed intervention kind of taking a stronger role.

SACKS: I love this chat, this question, Jessica, because I think oftentimes when China issues laws and regulations in the digital cyberspace, we often dismiss them as because of the political system that they came from. Right? China's authoritarian, so the laws don't matter. They don't have rule of law anyway. The government can do what it wants. But I think what's so interesting is that as my friend Kendra Schaefer from Trivium says, there's always a little bit of sweet and a little bit of sour in every Chinese law and regulation. And so when we look at, for example, the generative AI rules, which I should say is just one of a sort of matrix of a broader AI governance framework that we've seen unveiled over the last few years. And we can talk about what some of the other key rules have been. But if we look at it, we can see multiple conflicting goals at once, right? So on the we've I talk earlier about information control and security. So right up at the top of the regulation, you have a provision that says adhere to core values of socialism and that the content must not subvert state power, it must not incite succession. You know, this is sort of right up front up there. It's clear this is a Chinese law or regulation. But "yes, and" is the key. So if you were to read this regulation blind and didn't know it came from China, you would also see language that surfaces in debates in Washington, in Europe, about addressing the risks of digital harms, because you have multiple stakeholders within the Chinese political bureaucracy. You have academics and regulators who are also genuinely trying to address problems in AI that their counterparts in the West are. So there is extensive language around data protection, using data on the basis of consent, nondiscrimination, and non ex- - which is really driven by concerns around exploitation of delivery workers and sort of algorithmic harms, tagging mechanisms related to identifying content and a whole part about veracity of the data. So, you know, this is the sweet and sour of these regulations. And the question is how are they going to actually implement this thing when there are so many paradoxes within one regulation? But the key is China's moved out ahead. They've put a stake in the ground and we can't dismiss it outright because they're kind of a laboratory; they're a petri dish for a lot of the debates about AI governance that I think are happening around the world.

BRANDT: [Thanks, thanks so much. Marietje, just following up on that, you know, Europe, too, is out ahead on on regulation. And so I'd love to get a sense of what this looks like from your perspective. I mean, we watched through the winter and it looked like generative AI wasn't going to feature prominently in, you know, plans for regulating AI. And then all of a sudden in the spring, there was this flurry of activity to update, update those rules. And so I'm curious, like what changed? And also what do you see as the implications of the latest legislation? I think you're on mute.

SCHAAKE: Sorry. I apologize to everyone because my Zoom was having issues. I'm on my phone, which is like most uncomfortable for a great event like this. So on the AI Act, which I know many in the U.S. think about as the EU AI Act, it's important to know that it's still being negotiated. So it's not finalized, it's still being worked on. And that's also what you were referring to. So the initial drafting, which is typically done by the European Commission, as it was this time, started two years ago. And so this was supposed to be a comprehensive law to deal with artificial intelligence. And the decision was made to try to build as much as possible on existing policy instruments with regard to, for example, how AI applications impact people, and not so much looking at the technology in a more horizontal way, which could have been a very different starting point to any legislation. And we can talk more about what the tradeoffs are there or what the consequences are. But in any case, the European Commission decided that the application of AI comes with varying risks. So they basically create a spectrum of risk from high risk to low risk thinking of, for example, the consequences of an AI decision being that somebody loses their freedom if it identifies the risk of recidivism, for example, or that they lose their social benefits, if there is an AI enabled assessment of fraud, or selection of access to education or employment. You know, consequential decisions for people's everyday lives all the way down to much less consequential applications such as, for example, customer service. But I'm sure we can have a debate about that too. But, you know, provided that the customer service AI doesn't discriminate,

for example, which is already illegal in any case. This was sort of the gist of the law and there was debate about, you know, how much use of of AI for biometric identification and data gathering would be acceptable and should there be a carve out for law enforcement. And so there was a lot of politicking around what the law should look like in its final iteration. And then, indeed, as you referenced, generative AI kind of disrupted not only many sectors, but also the political debate in Europe. And so the European Parliament, which was the last of the three institutions to vote on its position vis-a vis-the EU AI Act, decided to include it in the sense of including foundation models and looking at how those can can lead to risk and looking at how the responsibilities of companies should be taken on board. And so I think there's broadly consensus that that is a sensible approach. But what you see is going from an applications-focused law to a much more sort of horizontal AI-focused law.

And I hope that the lesson learned for everyone who is now on these negotiating tables is that we may not know what will come next, but we do know that something else will come next. The generative AI is certainly not the last disruptive iteration of new AI application and that it's important that these laws are made in a way that they can incorporate these future uses and that, for example, and I think this is really important, enforcement is not only effective, which is a lesson learned after the General Data Protection Regulation, where enforcement is very fraught, but also that the enforcement agencies, the regulatory bodies, oversight bodies, watchdogs are equipped to really understand where the next wave might be coming from and that there is not going to be a challenge if it's not explicitly written into the law. So I imagine that these will be the types of discussions that the negotiators take on. There's also some aspects that actually, I believe have been overlooked, probably unintentionally, such as access to data for researchers, which has been a huge issue in the Digital Services Act, something that at Stanford, you know, we think about a lot. How can there be understanding of large language models in the public interest, not only steered and decided on by companies in the interest of their new products or competition with their with their lead competitors? But really, how can academics, but also journalists, civil society leaders, better understand how these models work? And so hopefully the negotiators will leave some room to to incorporate access for researchers. And so that just gives you gives you a little bit of a flavor. I'll mention one last thing, because we're talking about geopolitics, and I think it's something that may be overlooked by some people who're not looking at the at the details of this law much. The AI Act in the EU explicitly excludes military use of AI, and that actually puts it really in a different sort of corner compared to the Americans and the Chinese, where of course military uses of AI are actually incredibly important in driving considerations around national security investments, notions of competition, export controls, and where risk in a geopolitical sense comes from. And the reason is not because people didn't think it was important, but because the EU doesn't have the power or the competence, as Europeans call it, to deal with national security in the same way that the U.S. federal government has. So we will have to see how that plays out because of course every member state will will have to navigate the guardrails for military uses of AI or investments, and then somehow all those different decisions have to relate back to this overarching EU AI Act. So to make a long story short, it is going to be a comprehensive law that will definitely set the tone and the world is watching and there is anticipation that other countries will adopt parts or the general framework of the EU, but it is not covering military uses and there are still some some key issues to be worked out by the negotiators.

BRANDT: Thanks. That's very helpful context. Obviously, as a researcher, I'm biased and as are my colleagues, but I think we agree that researcher access is a critically important issue. As these negotiations are continuing, of course, there are conversations between the United States and Europe. There's been discussions over aligning investment controls, you know, of course, with an eye on China. And I was curious if you could just help us get a sense of the latest on that front. And also, like what of the TTC? What work will the TTC do, you know, on these issues?

SCHAAKE: Well, I think a lot depends on what the U.S. government is willing to do. I mean, it's pretty clear where the EU is going and I think they would be very interested in aligning with key allies such as the United States. But the political reality in the U.S. is obviously different, which explains why the Biden administration has doubled down on enforcing existing laws because

it's simply not to be expected that the Congress, as divided as it is, will come up with new laws. But of course, executive orders or other ways in which the Biden administration can act on its own should not be excluded. So I think one key question that is both important in the U.S. and in the EU and that has been discussed also in the G7, for example, is whether something should be done immediately. So yes, the EU is working on this AI act, but it will take, let's say two years before it will be fully entered into force because maybe six more months of negotiations and votes and then there's usually a transition period foreseen so that companies can adjust, governments can adjust, everybody can adjust to the sort of law that's coming. But of course, generative AI causing all kinds of questions and challenges and risks as well as we speak, so nobody really wants to wait two more years. And so there there's discussions about should there be some kind of voluntary agreements, should there be some kind of code of conduct, should this be done at the G7 level, transatlantically? Or is the U.S. going to going to be more interested in in doing something with their companies? I mean, we've seen Biden inviting the CEOs of some of the leading AI companies. Some people were surprised such a focus on on just corporate leaders. And then there were other working sessions with academics and civil society leaders. But I think essentially what the United States does is going to be at least if not as or maybe even more important as what the EU does, because the the power of these companies is so concentrated in the hands of a few U.S. players now. And so whether they will be subject to restrictions or oversight on the U.S. side, that is more than what we already know from existing law, I think will be a key question that Americans can answer better than Europeans.

BRANDT: And with that in mind, Chris, speaking of what will the United States do? You know, I'm curious for your thoughts on how, you know, the explosion of interest in this subject is shaping debates in the United States on these issues. And what do you sort of see as the trajectory of those debates and what of Schumer's, you know, safe innovation framework and all the rest?

MESEROLE: It's a great question. And I would just, I think, pick up some of the threads that Samm and Marietje kind of already laid out. I think there's really two conversations within the United States that are that are playing out in tandem. I think they've been separate so far, but they're about to converge in ways that I'm not sure has been fully anticipated by either side on the on the kind of pure AI governance side. So, you know, governing AI for, you know, commercial or consumer applications setting aside some of the strategic and national security implications. You know, I think there's some low hanging fruit that that, you know, Schumer and others and up on the Hill might be able to achieve that's very, you know, on par with what you might see in the EU or even in China. Things - I'm referring in particular things like disclosure, right, if there's a generative AI photo of someone in your political ad, like that should be disclosed that that was created by generative AI. I think, you know, there may be some kind of certification requirements that are that are kind of very low hanging fruit that might be able to get through as far as making sure that these systems, you know, have been tested or that there is some form of transparency around what kind of testing they've undergone. I think beyond that, it's going to be hard to get much through this Congress. And as far as you know, I don't anticipate seeing anything remotely like the EU AI Act and its scope or ambition coming through this Congress. And so I think on the domestic front, we'll see maybe a repeat a little bit of what happened with GDPR in the sense that a lot of the kind of data privacy protections that were implemented in the EU, you know, ultimately to some extent became de facto global standards even in areas where they, you know, outside of the EU, including in the U.S. I think with AI you might you might end up seeing something fairly similar there. Absenct, absent that, I think we're much more likely to see that different regulatory agencies we have in place lean into the regulatory regulatory authorities they already have and update them for an AI era in an AI economy, but I really don't see a massive kind of sweeping bill on par with the EU AI Act or even on par with what China has been doing with with its AI governance approach.

On the strategic side, I think this is where things are about to get pretty interesting. I think, there's two things that I think the audience should bear in mind when it comes to what's happened over the last couple of years with respect to the U.S. approach to China and technology in

particular. The two big kind of - I think there were two core assumptions to the way that the U.S. political establishment, let's say, kind of viewed China's tech development in the past. One was going all the way back to the early 1980s when China first started to kind of lean into this strategy of adopting technology and trying to move up the value chain. If you go all the way back to that era, the two assumptions were, one, that if China kind of, if this, you know, strategy succeeded and China kind of got a lot of market share at the low end of the tech value chain and then increasingly moved up that they would feel like they had more at stake in the international order and that they would do more to kind of uphold that order as they became more prosperous on the back of high tech development. The second kind of core assumption was really one that I think reflected a lot of complacency within within the U.S. and some of our allies, which is that we never I don't think anyone really expected China to catch up and to kind of start at the low end, but make it all the way to the striking distance of being able to produce leading-edge systems. And within two or three years of the Xi regime, I think both of those assumptions were, you know, kind of invalidated. Where there was a concern within D.C., I think at the very end of the second Obama administration and then through the Trump administration and this White House, that the Xi regime was not going to kind of be acting in the interests of the existing international order that we've had over the last 40 years, in fact, it might actually try to undermine it in ways that allow it to exercise greater authority around the world. And then the second assumption was invalidated when, you know, about five or six years ago with the prior version of deep learning, you know, there were some models, you know, some natural language kind of translation models and computer vision models, for example, that outperforms anything that was, you know, state of the art within the U.S. and Europe at the time. They've since fallen, you know, as as we said earlier, you know, they're now maybe kind of two or three years and just kind of coasting behind us. Once there was a new architecture for deep learning called the Transformer that they haven't kind of invested quite as highly in, one of the reasons being it's not as easy to align with what the creators of these systems want, wanted to do. But they're they're clearly within striking distance and they can kind of produce technology at par with us or better. And I think for a lot of national security strategists within D.C., they look at China and they say, you know, they potentially can kind of outcompete us on some of these technologies and we can no longer assume the status quo is that the United States and our allies are going to be able to outperform, like our technology will not necessarily outperform Chinese technology in the future. And in the military realm, that has pretty significant implications, especially for those in the Pentagon who are looking at, you know, Xi's declaration that -- not that he's going to go to war, I hope, I want to be really clear that he has not come out and said that, but he has said that he wants the PLA to be prepared to take Taiwan by 2027, which has really gotten the attention of a lot of folks in the Pentagon. And you couple that with their kind of tech ambitions, and that's produced this sense that the U.S., I think, needs to act now in a ways to kind of curtail, you know, advanced AI capabilities within China as much as possible.

The reason I bring all that up is that that effort to curtail China's tech development, it started with semiconductors, which is kind of the low-hanging fruit, right. That you start with by denying them access to the highest, most capable or most performant chips that are used to train large AI models, but you're not doing, you're not denying them semiconductors just for the sake of denying them semiconductors. You're denying them semiconductors to deny them the capability to train the kinds of models that you're afraid of, of them having. What's going to come out of that after the semiconductors is new restrictions on outbound investment and controls, right. Because it doesn't make sense to deny them semiconductors, but then allow Western firms to invest in Chinese semiconductor manufacturers and other AI startups. Beyond that, once that's locked down - and also, if that is your goal, is to deny the kind of capability, advanced capabilities to China, that also then puts you in the position of trying to lock down cloud computing clusters, using these chips outside of China and beginning to put in place kind of restrictions on the ability to rent cloud computing at time so that someone in China can't kind of rent a cluster outside of China to produce these capabilities. And then that kind of brings me to the last step, which is eventually, if they really want to deny access to certain like high-end capabilities by virtue of these models within China, they're going to have to lock down the models themselves, including even potentially open source versions of those models. We just saw Llama 2 released by Facebook yesterday. That's an extraordinarily performant and capable system that was open source. What this is going to do is if

the administration wants to deny access to advanced kind of reasoning capabilities within some of these systems, they're going to have to define what those capabilities are and impose mechanisms to assess these models ex ante, so before they're released, of what those capabilities are in the models, which puts them back and that puts the discussion back in the realm of the AI governance debates we've been having in the past in terms of in terms of model access by the government, which is something historically the U.S. in particular has been very averse from doing. But I don't see how now that we've started to go down this road a little bit, I don't know why you would kind of stop short of going further in that direction. And if they if they keep doing that, this this security and strategy conversation around technology is going to very much merge with a governance conversation. And I think nobody really knows yet within D.C. exactly where things are going to land. But and I'm not sure how many steps ahead this has been gamed out, but that that's kind of where I see these two trends within AI governance broadly in the kind of AI strategy debates playing out in the future.

BRANDT: Thanks, Chris. That's so helpful. I want to encourage folks to submit questions to events at brookings dot edu or with a tweet to hashtag geopolitics. And I am already getting a couple. But before we go to the some of those questions, Samm, I just would love to get your reaction. Is that like comport with your sense of the trajectory of the debates? Do you see pause points short of that sort of end point and and in particular, outbound investment screening? You know, if that's coming, what kind of impacts do you think that will have on AI development in China?

SACKS: Maybe I'll just talk about outbound investment a bit, because I think one of there's been a lot of we thought it was going to come and then it hasn't. And my understanding is there's been a lot of internal debate in the administration, in particular related to our conversation around how to scope key terms like AI. Of the sectors that have been identified, quantum semiconductors and AI, what does AI mean in the context of an outbound investment restriction? And I don't think there's been consensus on that. And I'd be curious to hear from Chris if you have any views if, as the governance conversation merges with the strategic conversation, maybe even from a prescriptive standpoint, if you have a view on what that definition should look like and any views on it, because I my sense is that's we are now, you know, Secretary Yellen from, from the the sidelines of the G20 this week, said that the scope of the outbound investment restrictions are going to be very narrow and targeted. And despite a lot of sort of bark from China hawks in Congress, I think there has actually been, it's going to be hard to issue an outbound investment restriction regime that is going to have a broad set of controls, because I think that there's going to be a lot of pushback to that. So my sense is it is going to be narrowly scoped. But the hinge point is on this definition of AI and Chris, be curious if you have any thoughts about that.

MESEROLE: I certainly have thoughts. I'm as curious as you are as far as like how they are actually going to define this, because if you're trying to do, implement some form of outbound investment screening on AI, you know how you define it, one way to define it, there's kind of two classical ways to define AI. One is to actually define it in terms of the technical capability or technical kind of specifications of a particular model, particular chip, etc.. Another is to say, at a little bit more of an abstract leve what you mean by in terms of the capabilities, like what can it actually do in terms of how well it can reason, how well it can respond to human queries, things like that. And, you know, I don't know, you know, in the end with semiconductors, they very much target it by technical specifications. I don't know that they're going to be able to do that with outbound investment screening because you don't know, you know, if you if you think about investing in a tech startup, you don't really know what the space is that they might, of what they might create. And certainly not with the kind of clarity you might need to be able to map out in advance exactly what kinds of capabilities you don't want these investment dollars going towards developing. And that kind of leaves them with, I think, the only option of trying to define it in terms of some abstract definition of what they mean by AI, but then the dual use nature of AI makes that also very difficult. I think if you were going to say, you know, it's a system that can kind of learn on its own and reason on its own at a certain level, how they would carve out architectures that have dual-use purposes from that restriction is going to be really challenging.

Just as an example of what I mean by that, if you look at one of the most important breakthroughs in my view, like positive breakthroughs in the AI space over the last five years is AlphaFold that DeepMind kind of pioneered and kind of allowed for protein folding. And it's going to be, you know, there's massive kind of positive benefits from being able to do that kind of scientific research. The problem is if you were if you were kind of screening somebody for in from like an investor in DeepMind that's doing biological research like that, the architecture that they used underlying AlphaFold could very easily be ported into domains to build AI systems that have very specific military or national security applications. So it's the architectures that underlie a lot of these models are agnostic to many different kind of industries, or they're kind of capable of being used on leverage across many industries. So I don't know how they would say prevents an outbound investor from investing in DeepMind for AlphaFold or for like a national security application of something that DeepMind's producing, but not for something like like AlphaFold, where I think they probably wouldn't want to kind of capture that in what they're doing. And so I don't know if they're going to define it in terms of just some general idea of what it is. It's going to be very hard for them to to very specifically target a, you know, abstract or conceptual definition that's capable of making those distinctions, you know, And that's not even to really get into the idea that, like startups often pivot over time and things like that. So it's my my sympathies lie with whoever is going to try and design this this kind of regulation, because it's not going to be easy for them to do.

BRANDT: Marietje, I'm curious whether you think Europe will get into the game and be willing to coordinate on this or are these definitional issues and some of the fuzziness here are going to get in the way?

SCHAAKE: Oh no. I think the EU is definitely willing to coordinate. And, you know, there is an active process also in coordination at the G-7 level, which, you know, is of a different, different nature, but it's politically relevant. And I think that that's something that we shouldn't forget. As I was listening to the questions of, you know, how important definitions are, of course they're extremely important. But there's also the political power, for example, that the Chinese Communist Party may use to single out certain applications just because it feels concerned about it, as it has done, you know, ad hoc, restricted corporate corporate developments, you know, as it as it was unfolding. And I think that room for maneuver is simply less and less likely to be used in Europe. There the letter of the law is really guiding or the mandates given to the enforcement agencies. And so I think we should we should keep appreciating the differences in the political systems and how they bear on what room corporations get, what values get baked into some of these AI applications. For example, is it considered a concern when disinformation flourishes, or are there those who see it more as an opportunity, just to name one example? So I think there is definitely a an appetite in Europe to collaborate with like-minded governments and to see that in a wide way, to also really consider Global South relations in looking to see whether there's a need for, you know, sharing knowledge or capacity in order for these countries to deal with what AI may mean in their local context and to see if there can be partnering there. So I really hope that that this is a spirit within which at least the like-minded democratic rule of law-based governments are going to work together because these developments are global to some extent. I think a lot of the harms will come from the least responsible actors. And so it's really going to be important that those who do feel like this technology should fit within the the much appreciated principles that the rule of law and open societies hold should really step up to make sure that those values are not swept away with, you know, disruption after disruption. So from what I can see, and I was in Brussels spending time with the European Commission yesterday, there was definitely a spirit of of seeking collaboration and partnership, not only transatlantically, but also globally.

BRANDT: You're speaking to a question that we received a handful of times from members of the audience who think there seems to be a sort of a collection of questions around international cooperation for AI governance and and some interests, specifically in areas that are ripe for cooperation. So this is you know, we've sort of mentioned one or two in this conversation, but I'm curious if you think that there are others that, you know, maybe haven't received quite as much attention, but potentially could or should.

SCHAAKE: What I do think it has received a lot of attention but has not been explicitly mentioned today, which is, of course, the relationship between AI and weapons systems, which I think is typically something that should be dealt with globally, and similarly is AI is part of a broader question of how international law, including laws of armed conflict, but also international humanitarian law, applies in the context of digital technologies and cyber. So, you know, in theory, a lot of international organizations and political leaders have stated and and made it official that law, as it applies offline, should apply online or should apply in the cyber context, too. But exactly how and what consequences should be, for example, with cyberattacks, you know, is there attribution, what should be proportionate retaliation, sanctions and so on? There is a lot of ambiguity, and I think that ambiguity at times is used, but it also leads to impunity. You know, the perpetrators of quite serious acts go unpunished. And so I think there are a lot of areas where, if that discussion had not been completed or led to sort of, you know, clarity in terms of how international law applies before generative AI broke through, that that discussion has not gotten any easier as a result. And, you know, overlaying or underpinning a lot of these questions is what is what is the the proportion of power that private companies should have in governing these global and very influential matters? And where should states step in and on the basis of what rules and principles? And I feel like that relationship question will become a part of many more international negotiations and initiatives, simply because the power concentration, the data concentration, the Compute concentration, the understanding has has only gotten more comprised to to, let's say, a handful of companies. And they have enormous discretion to make decisions about what they consider risky, what they want to do about it. But those decisions have consequences for how human rights are upheld, how conflicts are fought in the battlefield, what economic advantages are to be enjoyed by whom. So I think that sort of a layer over a lot of these discussions has only become become more important.

BRANDT: Did others want to come in on this?

MESEROLE: I can come in just briefly on a kind of foot stomp the importance of global governance for AI within the military space. And I think there there's a whole body of work that needs to be done around international humanitarian law in clarifying some of the ambiguity, there's a lot that needs to be done in terms of setting up communications channels and kind of confidence-building measures so that if there is a scenario that plays out, the uncertainty doesn't kind of lead to unintended, you know, either conflict onset or an escalation of conflict. I would also say that this isn't necessarily hypothetical. I mean, there was a, some of you may some of the audience may recall an attack on Abqaiq oil fields within Saudi Arabia a few years ago by drones that we ultimately figured out were from Iran. But in the moment, it actually wasn't clear where those drones had come from and how the strike had taken place. And so there was this period of uncertainty about the attribution of that attack that was made possible, in part by the virtue of the kinds of technologies that were, I think we'll look back at as fairly kind of early forms of, you know, ultimately what will be autonomous systems. And so I think we, hopefully that will serve as a wake up call of like, we need, you know, appropriate protocols in place to be able to manage those kinds of crises effectively. And the other thing I will kind of say on on this point is I think we desperately need kind of global norms around AI and its use within military applications in particular as kind of major regional powers who have the capability, like the advanced tech capabilities to on to build out very sophisticated weapons systems of their own, you know, start to sell those abroad. And I'm thinking of countries like Turkey, which just kind of announced a major deal with Saudi Arabia earlier this week around for some of its drones. They're not the only kind of major player in this space. There's Israel, there's Brazil, there's India, there's others. So it's not, even though the discussion around kind of military applications tends to focus on just the U.S. or Russia or China, it's actually a much bigger issue. And I think we we you know, I was I was heartened to see the you know, the State Department put out a political declaration that I think was designed to try and gather more kind of global attention to this issue. But I think that's something that really, you know, it needs to go way beyond just the State Department's effort there. I think I think there's a lot that needs to happen at a global level to coordinate around that effectively.

SACKS: For all the conversation around cooperation with allies and partners in democracies, I think something that I have not seen sort of any real robust discussion on has been how do we coexist with an authoritarian power that is using technology to stay in power and to monitor and do all kinds of things that we could have a whole discussion about: the human rights abuses from data enabled apps in Xinjiang, to what's happened in the Zero COVID policy, really disturbing uses of technology in China. And yet we have to coexist with this tech-enabled authoritarian power. And I think there's so much question, there's so much discussion around export controls and investment restrictions and how do we collaborate, collaborate with like-minded governments. But the core question to me is how do we coexist with China in this space? And I don't have the answer to that, but I think that that's a really hard and important conversation that has to be had. You know, I was recently with Chinese scholars and academics at a conference in a sort of neutral third country location and sitting around the table, you know, no government there. But I was talking I had two parents on either side of me, and we were talking about what it meant to raise children in the era of AI. And I just found it so refreshing and important to be able to have this conversation with Chinese scholars and practitioners. We all have young children. We're all very concerned about really similar issues, and I don't see a space for that kind of conversation in this political environment, and it's really tough and I just think it's really important.

SCHAAKE: Can I add one thing? I'm sorry.

BRANDT: No, go ahead.

SCHAAKE: Thank you. I agree. But I think here the task for the United States government is to actually make it much more clear what kind of model of regulation it believes in, because actually negotiating internationally or coming to any table, whether it's for dialogue or for treaty negotiations or for for anything else without more clarity on what the model looks like that you want to put up for discussion or dialogue is hard. And I think the price that the U.S.will pay for inaction in that sense, or for trusting the market or for choosing a liberal hands-off approach in international negotiations will actually become more clear because I agree completely that there has to be coexistence and hopefully clear terms and boundaries on the basis of which that can happen, that will be respected and so on and so forth. But I feel like one of the disadvantages of the U.S. not being more clear in terms of domestic rules about what it what it really seeks is that it's going to be harder to negotiate.

BRANDT: I'm curious if there are other things that you see, you know, Washington and Brussels could be doing differently to foster the kind of sort of environment that you are calling for. I mean, Marietje just offered one idea, but if others have thoughts, I welcome them.

SCHAAKE: Well, the difference is already quite significant between Washington and Brussels in the way in which China is perceived. I think the views are perhaps converging a little bit, but I don't know another city where a concern for China is such a leading topic as Washington, that is certainly not the same in Brussels for a variety of reasons. Of course, there is a war in Ukraine that really occupies people's political agendas for for understandable reasons. But even looking beyond that, I think there is already a difference. And, yeah, the way in which the U.S. and the EU and other partners work out those differences and how they each relate and how they collectively relate to China will also be extremely important. And I know that there is also quite a bit of frustration about the sort of forceful way in which the U.S. government is bringing on board countries. Take, for example, a company in the Netherlands called ASML that has not been able to export as it had intended. And so I don't know how long that political capital in the U.S. will last, certainly if there might be an administration change.

BRANDT: Chris, did you want to come in here?.

MESEROLE: Yeah, if I can come in just briefly on kind of areas of potential cooperation, I first of all, I agree with everything that Marietje laid out in terms of AI governance writ large. It's hard. You know, the U.S. needs to get its house in order a little bit on kind of how we want to

govern AI before we really start having any kind of earnest conversations with them about how to how to have or how to govern AI. I think there are two areas, though, that I think we might be able to focus on and have some forward progress which which are kind of areas of shared vulnerabilities, right. So there's there's one I think there's a room for a lot of cooperation around AI safety for things like, you know, autonomous vehicles, etc., where it's not know that, you know, both the U.S. and China, I think, are looking for solutions to make sure that some of these safety-critical applications of AI are, in fact, safe. But the probably the biggest area where I would love to see a more, you know, a concerted effort at a very high level and it's maybe somewhat paradoxical given how sensitive it is, but I actually I, I think the biggest opportunity for a breakthrough would be around nuclear command and control and insisting on AI oversight of nuclear command and control. If those haven't seen or tracked the U.N. Security Council briefing on or meeting on AI yesterday, Yi Jiang, who was one of the commentators from China who spoke there, and he explicitly mentioned in his comments that humans should always maintain and be responsible for final decisionmaking on the use of nuclear weapons, which I find it very hard to believe that he would say that kind of thing without it kind of having some form of authorization within the Xi regime for him to be able to do that in that forum. That is something that the U.S. has been pushing for and others have been pushing for, and not just the U.S, I mean, other other major stakeholders globally have been pushing for is for nuclear-armed countries to come out and say publicly that they do not intend to ever have to have final decisionmaking authority over nuclear command and control. It's a very specific issue, but it's a high risk one where I think all parties really do have a shared kind of concern that we get that right. And it seems to me that that might actually be an area where we might be able to make some progress in terms of AI governance.

SACKS: Chris, thank you so much for flagging that. I think that's really important. And I just wanted to to shape to share that in 2021 the an expert could be, in in China an expert committee on AI governance issued a series of sort of high-level guiding principles for ethical norms around AI. And in fact, that concept of humans maintaining control over AI and bearing ultimate responsibility for the systems is stated in those guidelines. So it is the it is sanctioned official Chinese policy. So maybe something absolutely worth worth exploring there.

BRANDT: Really helpful. I'm looking at the questions that we've been receiving from audience members, and there are a handful on a slightly different topic. So shifting gears for a second from cooperation to competition. You know, I think a number of questions are focused on how or whether, you know, regulation of AI in democratic societies will and in particular in the United States will slow down the United States or its or its partners in the competition with China. And I'm curious for, Chris, I think you may have a view on this.

MESEROLE: I have exceedingly strong views on this. I'll try to be brief, but I don't think that there's a I think it's a false dichotomy to assume that if we regulate AI effectively, that that's somehow going to slow down and throttle innovation. I mean, I think that there are you know, there are many ways in which we can regulate and govern AI effectively that still allow us to innovate and still allow us to kind of be competitive both economically and strategically, but would not and in turn would give us the kind of accountability and transparency that we would want these systems to have. So I think it's a bit of a you know, it's an argument that comes up often about why we wouldn't want to regulate or govern AI effectively. I think it's, you know, you know, very much a bad policy. And I think, in fact, I would say it's the other way around that if we are able to put forward trustworthy AI and really regulate and govern these technologies effectively, we will be better positioned to kind of recruit allies and partners globally and have a stronger case when China or other authoritarian regimes start making their pitch for their kind of governance model of AI. If we have a better one and a more compelling vision for how to use these technologies in a way that's safe and trustworthy, that is in the long-term strategic interests of the United States and other democracies around the world. And I think it's a hopefully that is the vision that we'll be able to pursue in the future.

BRANDT: Thanks. Another question we got is about key technologies that support AI, and this the questioner mentions like the graphics processor and is curious like how can we know?

How can we predict? How can we, you know, sort of forecast what kinds of technologies will be key to future AI developments and and how does that fit into our perspective on regulation?

MESEROLE: Well there. I think the important thing is to understand not just what technologies are going to be key to the development of AI, but which technologies are going to be key that are also not easily substitutable, right. And so there's that for which there's no kind of readily available substitute or that it's not easy to recreate. And so Marietje mentioned earlier a company called ASML, which produces an advanced photolithography machine that allows chip manufacturers to etch the circuitry onto it, onto a chip at very, very small scales using light. And that machine is extraordinarily complex. And I think it's something like 100,000 pieces within that machine. It's the physics that go into producing that are unbelievably complicated, and there's really no other company in the world that's capable of producing that right now. And so it's not always clear, kind of like what the far-term future technologies are that will be strategic. It is fairly clear in the near term that like anything that ASML does to build on the system they have is probably going to be something that, you know, other firms around the world are not able to replicate, certainly not within a short time frame just because they're so far out ahead in terms of how that particular really important piece of technology is being developed. And, you know, I think in the far future, you're looking ultimately as we kind of get close to, you know, Moore's Law is kind of starting to reach the limits of physics in terms of what it can achieve. And, you know, you'll need to kind of go into, you know, quantum or kind of photonics where you have kind of light based computer systems or circuits, rather. And there I think it's just too early to say what the critical technologies will be within those supply chain ecosystems, because those those technology stocks are not mature yet. So it's not really actually clear, even if you wanted to do an export control around quantum computing, kind of, you know, different components within that stack, it's really hard to know where to piece together, like how to piece together a good export control regime because it's the tech is so nascent and it's not even clear where we have an advantage and where we don't. So I think, you know, we can do it on a very short time horizon based on what the current state of the art is, but I think it's a bit of a fool's errand to try and project beyond that far into the future.

BRANDT: That's great. One question we just got was how can policymakers - this is sort of related - good policymakers overcome the challenge of the pace of development with the pacee of legislation and in particular related to international cooperation?

SCHAAKE: I could say something briefly about that, just building on what I said earlier, I think it's not a question of whether there will be new breakthroughs or unexpected developments coming from tech companies, but rather how and when. And so I think legislators should build in capabilities to be flexible. So that could mean, for example, empowering regulators. It could also mean, which is an element of the AI Act as it is drafted now, that there is a designated set of experts who can continue to observe, for example, new applications of AI in this case and identify them as being high risk, medium risk or low risk, and that way trigger some of the mitigating measures that are already foreseen in the law. but that may not have just been mentioned specifically for one application or the other. So I think, actually, given the ongoing developments in the tech field, it becomes even more important to be very clear on what values and principles are, are the laws are or the rules or the agreements internationally are supposed to safeguard. Because that way, if you're clear about those, then you can have the mandated authorities probe, whether whichever new iteration actually, you know, violates these principles or values or stays within those boundaries.

BRANDT: We just got two questions in a row on the challenges of myths and disinformation. The question is, can you dive more into that, into the challenges as AI improves deepfakes and how this can impact geopolitics? You know, they offer an example about a deepfake of a politician released the day before an election. I know, Chris, you've done some work on the risks of deepfakes in armed conflict. I hate to wrap up our conversation with a focus on risk, but but nevertheless, let me let me kick it to any of you who wants to take a crack at that question.

MESEROLE: So I have a somewhat counterintuitive sense of what the real risk here is. I don't actually think it's about, you know, like we saw deepfakes of both the Zelensky and Putin at the onset of the Ukrainian or the Russia's invasion of Ukraine. And, you know, they didn't cause kind of mass chaos or demoralization among their supporters. What they did do is kind of rally up a base of support among the folks who are trying to set a certain narrative around the conflict. I also think you'll see similar kinds of efforts within like political campaigns, etc.. We're already starting, like DeSantis, I think used a deepfake recently of Trump and Trump's audio. And I think those those efforts are not necessarily to try and get somebody to believe something that isn't true, but instead they kind of just buy into a broader narrative that's kind of providing momentum and fueling the entire kind of campaign or movement that they're trying to get going. I think the other thing I'll say is, part of the report that you're referring to is also a kind of a a wake-up call for democratic governments to make sure that the military and intelligence services who are using or responsible for information operations don't themselves use generative AI and deepfakes in their own operations in ways that has blowback within democratic societies. I think that, for democracies in particular, will be fundamentally crippling if it's kind of frequently the case or if it's seen to be the case that democracies themselves are producing deepfakes that are used in different forms of conflict around the world or for geopolitical purposes. It will discredit and kind of undermine the claims of democracies and the moral authority they have when they try and push back on these these efforts themselves. And so hopefully cooler heads will prevail, at least within kind of, you know, democratic governments, about how these things should be used. But that, to me is by the single biggest risk, is that it would end up undermining kind of democratic legitimacy and authority within really important geopolitical contexts.

BRANDT: I couldn't agree with you more. I think, you know, democracies are at risk of doing more harm to themselves than to their competitors of that kind of activity. So in any event, I think we are approaching the end of our hour. So let me just say thank you to our panelists for what was a very enriching discussion. And thanks to all of you in the audience who submitted such interesting questions and helped fuel the discussion. So thanks again for your time today, and I'm looking forward to continuing these conversations on the future. Thanks.

SCHAAKE: Thank you.

MESEROLE: Thank you.