THE BROOKINGS INSTITUTION

WEBINAR

RESPONSIBLE AI FROM PRINCIPLES TO PRACTICE

Washington, D.C.

Monday, January 31, 2022

PARTICIPANTS:

Opening Remarks:

JOHN ALLEN
President
The Brookings Institution

Panel:

MODERATOR: MELANIE SISSON
Fellow, Center for Security, Strategy, and Technology
The Brookings Institution

JARED DUNNMON
Technical Director, Artificial
Intelligence/Machine Learning, Defense Innovation Unit
U.S. Department of Defense

MICHAEL GIBSON
Deputy Head, Defense AI and Autonomy Unit
U.K. Ministry of Defense

HEATHER M. ROFF
Senior Research Scientist, Center for Naval Analysis
Nonresident Fellow, Artificial Intelligence and Emerging Technology Initiative
The Brookings Institution

MICHAEL STREET
Head of Innovation and Data Science
NCI Agency, NATO

* * * * *

PROCEEDINGS

MR. ALLEN: Good morning ladies and gentlemen, and good afternoon and good evening, depending on where you're joining us from around the globe. It's so wonderful to have you with us. I'm John Allen. I'm the president of the Brookings Institution. I want to welcome you to our event today, “Responsible AI from Principles to Practice.”

Advances in AI are accelerating at a rate faster than at any other time in human history. And this rate is creating new conveniences and capabilities in increasing efficiency and productivity that we could not have imagined. And this technology is really revolutionizing virtually every aspect and facet of society, including industries like healthcare, communications, supply chains and more. But most notably and for today, national security and defense.

While its use in modern militaries is still somewhat limited, the extensive range of possible AI-enabled capabilities is truly extraordinary. From machine learning used in predictive maintenance, which has great potential, frankly, and advanced logistics, to AI-driven drone swarms potentially of the future, AI is poised to impact virtually every aspect of 21st Century warfare to include accelerating warfare. Yet, while the vast potential of AI appears significant, its likely impact also presents some pretty serious policy challenges for those committed to its responsible application.

For, indeed, while AI and machine learning will almost certainly enhance our national security, it will also raise critical questions about the ethical, legal, and societal use of these same technologies. And, of course, this is all taking place at a moment when competitors and adversaries like China and Russia have no qualms about AI and its use to empower their respective authoritarian models at home and abroad. Now, the U.S., NATO, and the United Kingdom are at the forefront of addressing this challenge. All three have stood up defense research and innovation organizations that are developing and adopting principles for responsible AI and are beginning to implement new policies to guide the integration of AI-enabled technologies into their respective military's enterprises.

So, today, we are so fortunate to be able to engage in conversation with experts from each of these countries and organizations who will discuss how they're applying their principles into practice and what their goals are for the future. So, with that, let me introduce today's guests starting in no particular order. We have Heather Roff who served on the Defense Innovation Board as it defined some of the very principles that we'll discuss today and who I'm very proud to say, we are very proud is part of the Brookings family as a nonresident senior fellow. It's always great to have you with us and, Heather, it's wonderful to see you.

Next, we have Michael Street joining us from NATO where he is the chief of data science and AI within the Communications and Information Agency's Chief Technology Office. Welcome, Michael. We also have Michael Gibson, who is deputy head of the Defense AI and Autonomy Unit at the U.K. Ministry of Defense. Michael, it's wonderful that you could be with as well and thank you for joining us.

And last, but certainly not least, Jared Dunnmon, who joins us from the U.S. Defense Innovation Unit, where he led development of a set of guidelines to put the DOD's ethical principles of AI into practice. Jared, thank you for that contribution and thank you for joining us and engaging with us today.

But before I turn the floor over to Melanie Sisson, who is a fellow in our Strobe Talbot Center for Security, Strategy, and Technology, and who will be moderating today's discussion, I would like to give a quick reminder that we are very much live and on the record. And please feel free to submit your questions via email either to events@brookings.edu or on Twitter using #responsibleAI. So, with that, thank you all for joining us, all the panelists for making the effort to join us today from around the world. You're going to make this panel very important and I can't wait to hear this discussion. So, Melanie, the floor is yours.

MS. SISSON: Great, thank you very much, John. I would like to start by simply reinforcing the warm welcome that John has provided. We're delighted to have all of our panelists and very grateful for their participation and we're also delighted and grateful to have so many of you joining us to learn from them today. And as John said, please do submit your questions through the available mechanisms.

As John noted, one of the things that challenges us as we think about the role of AI-enabled technologies in military affairs is just how many potential applications there are. AI already has changed how militaries engage in conflict through the use of advanced sensing, precision capabilities, entry platforms, and so forth and it's certainly going to continue to do so. It's also changing how militaries operate outside of conflict. How they repair equipment, manage talent, train service members and staff, communicate, make decisions, and more. AI and ML enable all of these tasks by increasing our ability to capture more and better data to process it, manage it, analyze it, and then to take action on that basis.

Although for many people, the phrase responsible military AI might bring to mind its role in the conduct of war. And perhaps even more specifically, the ongoing and very important international conversation about lethal autonomous weapons. The breath of uses of AI and ML means that the ethical responsibility of militaries extends well beyond these issues. Defense organizations also are responsible for managing appropriately the effects AI-enabled technologies have on the daily experience and well-being of service members and civil servants, and on the ability of citizens to trust that their defense organizations are providing security in ways that are compatible with their priorities and their values.

As John mentioned, this is an incredibly important element of work in this area right now as we think about international competition and expressing the things that are important about free, liberal, and open societies in our practices. So, while our conversation today might touch on laws, the intent here really is to have a much broader conversation about how to embed responsibility in the development and deployment of AI applications across the entirety of the defense enterprise.

Each and all of our panelists has taken up that question directly and are working to fulfill it within their respective organizations and in their research and in their participation in the public debate. We're especially fortunate, therefore, to have multilateral representation today. I think it makes this a unique opportunity to learn more about how our partners are thinking about responsible AI, how they're making progress toward putting their principles into practice, and about how all of these efforts can be really mutually supportive.

So, I would like to begin by turning to Michael Gibson who is joining us from the U.K. MOD. Michael, if you wouldn't mind starting us off by telling us a bit about the Defense AI and Autonomy Unit and where the MOD is in the process of developing its approach to defense AI and safe and responsible use.

MR. GIBSON: Of course. And thank you for the introduction. A pleasure to be here. It's very timely with where we are in our journey to understanding how these technologies can be used. And it's great to be on such an impressive panel. It's an exciting time for us. We are on the verge in the U.K. of publishing our AI strategy. That should be out in the coming month or so. That will be our first AI strategy. And I'll talk a little bit about how we have understood the various challenges, opportunities, implications of these technologies.

But if I just start first with explaining a little bit about the Defense AI and Autonomy Unit, where I sit. It's a function in the Ministry of Defense head office. We report jointly to our core strategy team and to also to answer to technology and enterprise. So, we've got an access to a huge volume of talent as well as, obviously, the key leaders within the department. And we set high-level policy and strategy. So, I'm very much a policy person and we seek to drive coherence across the enterprise.

And we've developed some really quite ambitious plans for defense adoption of these technologies. But it's not a Greenfield site. This work is already processing. We have something like 200 declared projects already underway across the department. Admittedly, some of these will be in different styles of the technologies, machine learning, advanced computational statistics, et cetera. So, AI tends to get badged to lots of different things.

But the purpose of our unit is to really ensure that we can make best of the technologies it matures. We recognize that certain use cases are controversial. So, we're responsible, again, for understanding what the right governance policy assurance mechanisms are in order to deliver the appropriate levels of accountability. And that's important on many levels. You touched on some of them. We have to retain the trust of our people within the department. We have to keep confidence of industry and the working with responsible actors. And we have to keep the license to operate from the public, absolutely critical.

And you also mentioned as did John, it's important for our global influence. So, projecting our values as a counter to what we are seeing as increasing techno authoritarianism. So, this concept of how we adopt the technologies of being demonstrably safe, responsible, and ethical is absolutely critical.

So, our approach to ethics. I mean, again, it's not a Greenfield site. Military ethics is obviously not new. We've got longstanding traditions in all of our countries and culture embedded across the force. But we understand that AI poses some new and different challenges. That's obviously a ubiquitous technology. You've got the issue of nondeterminative outcomes, of systems that learn and change throughout their lifecycle. You've got questions about predictability, the nature of human control, algorithmic bias, all of those sorts of issues.

So, there's a lot for us to grapple with. I'm very, very grateful for being part of this panel to exchange learnings. Within defense, we've spent 18 months doing our due diligence, I guess, you'd say. We've consulted very, very broadly in terms of military, academics, ethicists, technologists, civil society, industry. We've been learning from our partners. As I say, we're all on the same journey. And common approach is going to be critical to interoperability going forward. Obviously, that includes the U.S. and NATO. But also, I mention here the AI Partnership for Defense that's now, I believe, has 16 or 17 different nations involved in it. It's going to be a critical forum to getting that common understanding.

As far as our ethical approaches are concerned, I do have to put a caveat on my words. We're, obviously, prepublication, so I can't steal minister sandwiches by telling you too much about the detail of our approaches. But I would say you could expect very significant alignment with those principles that are being published by the U.S. and by NATO. We take very much a capability focus. We recognize and we published our integrated review in Defense and Security last year. We recognize this era of systemic competition, the need to deliver the best capability of our forces.

So, our approach has to be ambitious in terms of the effects that we seek to deliver. It has to be designed to enable not constrain. But, of course, as we're all here for, the golden thread underneath that is to ensure that our approach embodies our commitment to deliver AI systems safely and responsibly in line with our values and our legal obligations, national and international.

So, on the cusp of publishing, some thoughts quickly on implementation, how we're going to go about embedding this across the department and our wider defense ecosystem. Because it's important not to forget our industry partners there. I pick out a handful of things. I think the ubiquity of AI is a key challenge. It's going to impact everyone across defense from the S&T early stage, the capability development, to regulators, to the senior users who need to understand the different implications and choices and tradeoffs and challenges. All the way down to our end users. And that's going to require different levels of understanding, different training requirements from your senior officer corps through to your experts, from our industry partners to the classic professions, our lawyers, our commercial advisors, our other specialists. So that ubiquity is a big challenge.

Our approach has to be risk based. So, we're developing thresholds where we can understand, you know, where does heightened scrutiny and assurance needs to be applied to most value. We've got to mesh with existing processes. We can't just set up cottage industries to manage this stuff to separate governance processes, et cetera. It all has to be mainstreamed, built into our capability development. Built into our acquisition, our safety processes, our operational doctrines. And we have to recognize that commands have different cultures and structures. We have to find ways to apply things to the right context.

A couple of final points. We have to have appropriate levels of transparency about what we're doing. Yes, we need to respect security. But that transparency is going to be absolutely crucial to retain our license to operate. The right levels of independent challenge. I'd also say our understanding of these issues is evolving. And we're really in the foothills at the moment. So, this is the right time to start putting in place your principles, putting in place your frameworks so that we can build up those technical policy and ethical muscles, if you like, over time. So we can learn from the early and less contentious use cases so that as the technology matures, we're actually ready to start tackling some of those much more difficult challenges that you indicated early on. So, that's my introduction, how I'm starting to understand and think through the issues. But a big program of work ahead.

MS. SISSON: Indeed, thank you, Michael. I appreciate that. You've raised a lot of issues and I hope we can return to many of them as we go along. We'll turn now to the other Michael. And, Michael Street, if you could follow the other Michael, Michael Gibson's excellent example and share with us a bit about your organization. NATO did release an AI strategy last fall.

MR. STREET: Mm-hmm.

MS. SISSON: And so, I know the audience here would be interested in hearing as much about that as you are able to share with us in terms of the developmental process, any of the thematic elements that Michael raised that resonate. And certainly, about how you're moving with those principles into implementation.

MR. STREET: Okay. Thank you. And certainly, there's a lot of commonality with everything that Michael has just said from his perspective on the other side of the North Sea from where we are.

Just maybe a little introduction. I work in the NATO Communication and Information Agency. This is an amalgamation 10 years ago of a raft of basically entities in NATO that dealt with digital technology. So, we're now, we're effectively NATO's digital agency. So, we do everything from R&D through to service delivery and service delivery from computers on desks in headquarters to laptops on decks of ships or in theatre. And we have a sister agency based in Luxembourg that does all the nondigital things. So, within that we're very much on the execution side of development and execution side, as well as the procurement. So, the majority of NATO's engagement with industry goes through one of the two agencies.

You mentioned NATO's development of an AI strategy and we were a part of developing that. So, bringing up our experience of actually implementing machine learning projects and AI projects in a defense environment and often in a classified environment. And bringing that, sort of as our colleagues on the policy side of NATO HQ were developing that strategy, we sort of kept it linked to the reality of what was possible and also what was achievable from the technology side. So, we worked very closely with them from that perspective.

Now, I know that there is a version of the NATO AI strategy will be released at an unclassified level and made public and particularly focused around things like the principles for responsible use of AI. I don't think it will surprise anyone some of the principles that are being mentioned in there such as ethical, the legal use of data, nondiscrimination, things like that.

From our perspective, coming from the technology side, and I'm privileged to lead a small team of exceptionally talented data scientists and data engineers who work inside NATO developing solutions for the use by NATO. So, one of the aspects that's high on our mind is how do we prove that all of our projects are compliant with those principles of responsible use. Because it's very nice to have these principles listed but how could be tangibly say just as we will for the performance or the accuracy or functionality, how can we prove that we meet those requirements. And we're currently refining some of the processes that we've been doing over the past couple of years to sort of formulate to just capture them in a fairly, I don't want to say too -- but a rigid process, but to capture them in a methodical way so that we've got tools and techniques and mechanisms that we can ensure that all of these processes that all of these -- sorry -- all of these pilot AI projects that NATO is developing, we can look back at the principles of responsible use and confirm, yeah, we have a tick for all of these because we've been through a rigorous process for each of them.

In terms of the scope for what we do in our team, our agency supports all of NATO from the political level at NATO headquarters, the NATO command structure, the NATO force structure, which goes down into the nations. So, we have luckily to work with a very broad cross section of users across NATO. And we've been working for a number of years, yeah, since the last decade at least in terms of delivering projects to help users by giving them data science tools and AI-related tools that can help them in their day jobs.

The watch word of our team is better decisions faster, but it's through the use of technology. So, it's not applying AI for the sake of it, but applying AI because it's the way that we can get, we can help somebody, we can help analysts or commanders pull in this vast amount of data which they now have to deal with. Give them some guidance and some consolidation. Have models which will be able to, if not take decisions, at least make recommendations or help them to sense of the wealth of data that's being presented to them.

And in terms of the decision cycle, sometimes in some cases the decision cycle is extremely short in the operational side. But sometimes actually, the decision cycle might be developing tools that'll help people to do in a matter of weeks what used to take them a number of months. So, it's not always about, you know, operational tools which will instantly give some sort of recommendation.

You've alluded in the introduction to the use of autonomous weapon systems and some of those challenges. And I guess here we're quite fortunate to have a bit of a get out because NATO is a collective defense organization. NATO doesn't have attack capabilities, but it does coordinate the attack capabilities of its member nations. So, in that respect, we have a nice bounding in terms of the work that we're doing in our group that this is purely to support collective defense activities, not around offensive capabilities.

MS. SISSON: Great, thank you. That's a really nice introduction and explanation of the work that you do.

MR. STREET: Yes.

MS. SISSON: And similarly, I promise to do a couple of go backs on a few threads that you introduced there. Let's move first though over across the ocean to the other ocean to Jared Dunnmon who's joining us today from DIU. Jared, thanks for being here. You know, DIU recently released those AI guidelines, ethical AI guidelines for responsible use. And they're very thorough going and my understanding is that they're designed to help all of the participants in the process of acquiring these AI and ML technologies and capabilities to do so in ways that are really consistent with those ethical principles that the department has adopted. I'm hoping that you can share a little bit about both DIU first. It's a pretty unique organization I think within DOD. And so, if you can share with the audience a bit about what it is that DIU does and how it does it. And then talk us through those principles, the intent behind them and how they are being put into practice.

MR. DUNNMON: Sure thanks, Melanie. DIU to start with, the Defense Innovation Unit, is a fastmoving DOD organization that is focused on contracting with commercial companies to solve national security problems. So, if you think about it in that context, there's really three goals as an organization that we try to aim for. Which is, you know, accelerating DOD adoption of commercial technology, transforming military capabilities and capacities, and then also strengthening the national security innovation bases. So, the companies that the DOD works with.

In that context, what we do, and it's important to understand this because the guidelines that we released were very much in the context of the types of programs that we run, what we focus on doing is very much an acquisition process that is focused matching a commercial solution with a DOD problem. And so, what we start with from a very -- as a very brief example, is we start by curating a problem along with our DOD partners. So, we receive it, understand it, evaluate it, and then we confirm that there's actually a commercial market that can address that problem. And we want to aim in places where it's not when say a commercial solution, often, you know, research and development land is something that needs to be built that doesn't exist yet. It's not something that you can just buy off the shelf necessarily. It's usually in that middle ground where say it's early in a series A or B company. It's in a different industry that needs to be adapted for DOD use. Or maybe it's even at a large company, but it's an internal tool that hasn't been productized yet because the market signal hasn't gotten there yet.

And so, once we confirm that that's the case, what we generally do is solicit proposals to a problem statement. It's about two pages long. So, it's pretty short. We evaluate those proposals and get bidders in to pitch. And then we select a contract awardee and negotiate an agreement. And that process between posting that solicitation and getting that contract hammered out, you know, our goal there is to get that done in 90 days. So, we do that quickly.

And then what we then do is usually execute that protype project for about one to two years. And over the course of that project, if there is a -- if the vendor is able to meet the success criteria of the DOD partner, what we can do is write a success memo that says, yes, this thing did work, you know, this experiment worked. And that authorizes the DOD to then go acquire that at scale.

So, that's mechanistically what we do. We have six different portfolios that reflect areas where private sector investment has been enormous. Particularly with respect to public sector investment over the last, you know, relatively short amount of time. And those are advanced energy, AI machine learning, where I sit, autonomy, cyber, human systems, and space. So, I'll talk about this in the context of your AI portfolio, where I sit. It's important to note that a lot of the interesting work falls in the cracks here. And so, you end up often having projects that stand, you know, as was mentioned earlier, you know, a wide variety of different applications because usually what we're doing with, you know, AI systems, you know, I'd use that term broadly, is trying to increase speed, scale, or performance on tasks that are otherwise difficult for or inaccessible to humans.

And those can be things spanning, you know, knowledge graph construction, where you've got millions or billions of documents you want to understand the relationships between different entities in those things for things like understanding our supply chain. You have humanitarian assistance disaster response type of applications where you need to analyze imagery very, very quickly to understand how to best allocate your resources. Or even things like back-office business functions where you have folks who spend a long time, you know, performing specific tasks that we can accelerate with AI action learning.

So, in that context, you know, why did we start developing these guidelines? Well, really DOD, you know, put in a lot of work to say what are the ethical guidelines we're going to use? What are the ethical principles we're going to use and define how we build AI systems. And for those who are familiar, you know, in five words those were reliable, governable, traceable, equitable, and responsible. And those were released around February 2020.

Now, immediately when that happened, given the type of program that DIU runs when we're working with both vendors and government partners, we got a bunch of questions from our commercial partners just saying, hey, this is great, but what does this mean I actually have to do? So, concrete, brass tacks, how do I make sure that I'm aligning with these things? And so, we went through the process at DIU of trying to say, all right, let's at least for our programs, for the ones that we're running, not go and set policy, but bring out some -- put out some guidelines to say, you know, on these types of programs, how can we make sure that we're building and developing AI systems in a way that aligns with those ethical principles the department is committed to.

And so, the purpose here was really to make sure that those AI ethical principles are integrated into planning, development, and deployment of all of our DIU programs and prototypes to enable DIU stakeholders to effectively examine, test, and validate. That all of those program and prototypes meet those DOD ethical guidelines. And then, finally, to establish a process that's reliable, it's replicable, and it's scalable across our programs. Because you can do something that's bespoke for a particular program, but then trying to apply it to the, you know, both the, again, number of applications and just the sheer number of programs that could exist is something that we want to make sure is viable.

So, in brief, I'll walk through kind of what the outline of what those look like and then I'll stop there. And we can sure come back to some of it. But broadly speaking, what we do is we breakup the process of developing an AI system into three stages. So, planning, development, and deployment.

And so, in the planning stage, we ask, you know, we have really three lines of inquiry. You know, so, planning, development, and deployment. And within each of those lines of inquiry, we have a relatively linear process about, you know, kind of five overarching questions. And we tried to make it things that are kind of understandable across three different axes. You know, kind of folks who are junior to senior. Folks who are, you know, technical and nontechnical. And what we ended up doing was, you know, for in planning, we start by walking through these processes as follows. Things like, you know, first question. Have you clearly defined the task that you're doing? The quantitative metric that you're using to measure performance. And a benchmark to evaluate capability performance as well.

So, again, not just how are you measuring this, but how are you currently doing it today? Because most AI systems are built in a relative context, not an absolute context. If something is being done terribly today, an AI system that looks on its face to be mediocre, but actually provide a lot of value. Or in some cases, if you do the task really well already, you have to do really, really well for an AI system to add value.

Secondly, have you evaluated ownership, access, providence of it, relevance, candidate data? You know, third, are end users, stakeholders, and responsible mission owners identified? And fourth, have you conducted harms modelling to assess the likelihood and magnitude of harm? You know, fifth, have you identified processes for system rollback, error identification, and correction?

And so, again, you know, I'll go through these -- go through them very quickly. But in development, we focus on things like, you know, preventing intentional or unintentional manipulation of data. So, in development, we're actually building the system at this point. In planning, we haven't started yet. You know, next, have we defined a procedure for a reporting process versus performance opposed to point monitoring? Have we designated roles and persons with power to make and certify necessary changes to a capability? And then making sure that we're set up for third-party auditing.

And finally, to the last part, and the most important part and I'd say the part that really sets AI systems in many ways apart from kind of traditional software systems, is there's a constant deployment cycle. And this one is not linear. This is like a circular process. It's, you know, are you constantly doing things like task and data validation, harm assessment and quality control and functional testing. Because you usually can't hit the entire test surface of the systems that we're talking about here.

And so, to be able to have some confidence that you're deploying this in a way that makes sense, in a way that you can be confident these systems are going to continue to provide the value that you want them to provide, you've got to have an infrastructure for constantly doing these assessments and constantly, you know, checking your assumptions, et cetera. And that's something that in software we're often not necessarily used to doing outside of things like, you know, continuous integration, continuous deployment. And making sure that analogous infrastructures are set up for AI systems is a really important part of this.

So, I'll just end by saying that we deployed this on a number of different programs over the last year, year and a half. I can go into some of those a little bit later. But we've also published supporting materials here. So, the worksheets that we used that have some pretty detailed commentary are publicly domain. There's a workshop that we're running for training on how to use these things. We've got a white paper that describes, you know, those guidelines, where they came from, why they are how they are, a couple of case studies as well, and some lessons learned.

And some of the most important piece I'll say here is that we did that a) because we want to make it useful for as many folks as we can, but b) because, you know, this is our best working hypothesis. This is not saying like, yeah, we did this and it's done. There's going to be a consistent, you know, need for feedback, for iteration, for making sure that we continue to think about how we can do this in an optimal way. And I'm sure we'll always be working on our kind of best hypothesis. We'll never have it, you know, 100 percent right. But that's because the technology moves at a very fast pace and we need to be cognizant of that.

So, that's one of the things that I want to call out is that this is the result of not just, again, government going and saying this is the way it shall be done, but in reality, what this should reflect is a process where you've got private sector stakeholders, folks from academia, industry, government as well, bringing those different perspectives together and trying to align at a set of languages, set of processes that can be communicated across those different sectors. And so, we're excited to continue working on that. So, I'll stop there. I'm happy to dig in more later. But thanks for everybody for joining. I really appreciate being on the panel. Looking forward to more conversation.

MS. SISSON: Great, thank you, Jared, very much for that. That is a nontrivial amount of work I'm sure that you just sort of wrapped up into a mere moment's worth of comments. So, we'll make sure that we can sort of probe at that a little more as we go. First, before we get to some of that, I want to bring in Heather. You come from a bit of a different perspective on this, I think, having been part of the process of developing the DOD's ethical principles. And I'm curious for you to share as much about that process sort of in reflection here based on the comments of the other panelists.

In particular, I'm curious about the extent to which some of the elements they've discussed were anticipated in the process of developing those principles. And sort of what you see here that's very consistent with the conversations that were had and if there are things that surprise you.

MS. ROFF: Thanks, Melanie. And thanks everybody for inviting me and for tuning in. So, I wanted to just make note today that my comments are in a private capacity.

They do not represent, you know, the DOD or the Defense Innovation Board, or my current employer at the Center for Naval Analyses. So, what I'm going to say here is personal opinions. However, each one of these entities has a lot of great people working on a lot of these issues and I'm happy to discuss.

As far as what has been said before by the Michaels and Jared, you know, I would say that a lot of what they're doing was anticipated by the board and by myself. I was a special governmental expert attached to the board to do the primary authorship of these. Our process was stakeholder engagement. It was a lot of work engaging with folks within the Pentagon as well as with some of our partners and allies making sure that we had coordinated with them on what was going on. As well as engaging with academia and industry in a series of roundtables as well as public commentary.

And so, bringing all of that together and trying to understand what was really new about AI in this space and what was really new that the DOD needed to come out and set forth a series of principles to address the AI question, knowing full well that the DOD has been engaging with AI since the late '70s, with expert systems and things of this nature. So, how to be holistic about it. So, in doing so, we did anticipate the notion of interoperability. How we would communicate and how we would think about our allies and partners in what we would hope would be a norm cascade going forward. And so, it's lovely to see that that, in fact, is happening.

In terms of DIU taking it up is even greater, in my opinion, because I think the work that Jared and his team has done does a really a good homage to the work that I did. So, for those of you who are not aware, and it's not just the principles, that the board, you know, there was an S&P subcommittee that I did this work. There was a small paper and then there was a very large paper, a 70–80-page paper that probably caused all the grey hair in my head right now, that I had to put out. And a lot of the kind of nitty gritty of what that really means in terms of what continuous testing looks like. What we mean by assurance. What we mean by governability. What we mean by responsibility. Things of that nature. Are more even though it's merely 80 pages, in my opinion, it's also 80 pages. But I think in terms of operationalizing that, Jared and his team has done a really good job of putting kind of more meat on that bone and giving folks a way of thinking about how to take those principles into practice. And so, in that respect, I'm really happy to see the work that's being done.

I would also say that the only -- and the lines of effort, as well, in that that Jared has discussed were also quite anticipated in some of that work, which is great. You know, clearly defining your task. Clearly defining your benchmarks. Clearly defining your metrics. Understanding who your stakeholders are. Doing a risk assessment. I mean, risk assessments just generally are actually quite hard with these probabilistic systems that can be embedded in larger and larger context and context awareness and understanding, you know, the fact that these systems can fail spectacularly. And so, actually even understanding what the risk of failure looks like and failure ontologies is incredibly difficult.

So, you know, those harms IDs. Those risk assessments are really, really important. And they have to be done with kind of a grain of salt, too, right? You're making a risk assessment with a set of assumptions. And if your assumptions don't hold, then your risk assessment is going to be shifting one way or the other. And so, that error identification or that failure identification for the type of system, you really need to be quite careful about and understanding, you know, not just the platform or the system, but the system of systems, and the system engineering approach, as well as the human systems integration. Because that is something that I think gets widely overlooked is that in most DOD acquisitions, not on the protype DIU side, but on the big acquisitions programs, you know, we have had in the DOD at least an acquisitions reform to the 5,000 didact series. We have new software acquisitions pathways. We have middle tier acquisitions.

But in human systems integration is a key component of that acquisitions process, but oftentimes, is I don't think given pride enough place in thinking about how the way in which we integrate, you know, human systems to these other types of machine learning systems or AI systems is done adequately. Most often because we don't lack -- or I should say, we lack a lot of really good data when it comes to human cognition. When it becomes to human failure rates with some of these systems. Especially with talking about adaptive AI systems or personalization of these systems. Or even like not necessarily personalization, but even if you're creating categories of buckets of users and how that might look.

So, we lack a lot of data. And when it comes to that human machine teaming element, we don't understand that very well either. We lack benchmarks. There's a lot of things there that I think the human systems integration side needs to put into play as we move forward with that grain of salt of the risk assessment.

So, I'm really happy to see a lot of the work going forward. What I would say that one of the things that the project now kind of taking off my DIB hat and putting on my personal hat a little bit more is, you know, many of the members on the panel today have already kind of highlighted or flown by a few of these points. But I can't stress them enough that, you know, organization in culture is so important going forward. And when you think about the size and complexity of these organizations, you know, the DOD alone is the largest employer in the United States, right? It is by far the largest defense entity in the world. And so, when you think about the myriad number of people and incented structures and cultures within cultures, there's a lot of different things that can happen.

And so, while DIU's mission, for instance, I'll just pick on Jared some more, right, might be to do a rapid acquisitions in a prototyping to get that middle tier pushed out as a proof of concept. What might end up happening in a, you know, kind of perverse incentive way is saying, you know, here's widget, here's problem, here's a solution. Deploy, deploy, deploy. Without then thinking about not just the scalability of production and deployment, but the scalability of training. The scalability of trust acquisition. The scalability of the threats of how different missions might be amended or perceived or changed or workload flows.

And so, when we're thinking about kind of the scale within just how something is used, not just how something is acquired, it gets really murky really quickly. And then you kind of end of with his push, this incentive push to acquire these solutions quicker and quicker without thinking holistically or strategically about how you will integrate that into your force. And you end up with sort of a tyranny of small decisions, right? So, tactically you might think you're making really great decisions. But strategically, you might be hindering the way in which your forces operate if you're not thinking about that concept of operations at a better scale. Or at least kind of how you're going to be implementing that in that concept of operations knowing full well that there's going to be interoperability issues with your allies and partners. That your communications networks may or may not actually talk to one another. That you're going to have to interface with legacy systems. That you're going to have to have, you know, sticky cultures and bureaucracies, and, you know, gnarly, crusty old sailors who don't want to do things. And, you know, hot shot fighter pilots who want to do everything else, right?

So, you have to really think about all of these organizations and cultures as you're going forward as well. As well as your strategic vision of how you're going to fight and how these situations and these technologies, even from back office to this pointing of stick, really do put pressure on that culture. Because cultures and organizations are very, very sticky and they don't like to change.

And so, I would say that's a big one that I don’t hear a lot about. I mean, we do hear a lot about the acquisitions process and how it's a nightmare. And the federal acquisitions, you know, regulations in the United States are horrible. We have to go to other transactions authorities and all this other stuff. But the people side of it we say that people are our best asset. But it's also one of those things where we have to figure out the ways in which we incentivize people to adopt technologies and we incentivize the organizations to change accordingly.

So, that said, I will also say that there are, right, within we do kind of also a fly by on governance structures around emerging technologies and AI in particular. And I think, you know, as my distinguished panelists have alluded to, right, there's a lot of different ways in which we can govern technologies through our militaries, right? We can have policies. We have international legally binding treaties, right? We have international law if we're thinking about the use of force. In the DOD's case and likewise, you know, the MOD's case, right, we have issuances, right? They might not rise to the level of a policy, but they may be a directive or they may be some form of guidance.

And all of those issuances, policies, those laws, right, they have differing kind of weights. And they apply to different roles. They apply to different authorities. Those have to be coordinated. And often, they're not well coordinated. Or even if they are well coordinated, the responsible parties may not fully understand what that means for them.

And this is an important point because when we're thinking about, you know, what the commander might need to know when something is deployed into her hands, she's relying on the fact that the system has worked to get her a particular capability into her hands. Once that capability's into her hands, she's ostensibly supposed to have the training and the understanding of how to use, when to use, the risk profiles of that system, and to make appropriate judgments as to the use of any particular system. That's a lot of assumptions to build in there that the commander has the appropriate understanding of that system. When in fact, most of the time we're doing prototypes in this kind of whack-a-mole way and how it integrates with other systems. And so, I would say that the governance structures need to more coherently kind of jive.

And that's just the governance structure. It's not even thinking about the technical standards, which we are sort of lacking. The ability to write requirements well. And then the T&E and B&B assets that are required at least within our own, you know, the United States' own systems, right? So, we have the Test Earnings Management Center. But even within each of the services, right, actually coordinating, you know, does the navy's COMOPTEVFOR and marine corps' MCOTEA and all of these things. Do they have the people? Do they have the test ranges? Do they have the requirement? Do they have the knowledge? This is not a one-time problem if I can give you a widget or a tool and you've actually done your AI ethics principles operationally. It is a living, breathing thing that every single person has to think about and reflect about in their own day-to-day. And then what the downstream effect is, right? It's kind of a black mirror moment, I think, for the DOD going forward.

So, I would just say again, you know, there's a lot of things that need unpacking, but I'm really happy to see some of that coming forward today. Thank you.

MS. SISSON: Great, thank you, Heather, very much. It's wonderful to hear that you see as much alignment as you do and that you're heartened by the work that these folks and their teams and others have been doing to really embody those ethical principles in action.

So, I want to actually pick up on something that each of you has made reference to in your comments so far. And let's talk a little bit about this notion of interoperability. I'd like to ask each of you to talk a little bit about when you refer to interoperability and how your organization is thinking of it or acting upon it. What does that mean? And at what levels is that really necessary? Are we looking for consonance? Are we looking for deep technical alignment? What are we looking for? And how do we start moving in those directions, at least as it seems to you from where you sit today? So, I'll actually I'll pop back to you, Michael Street, first and then maybe we can go to the other Michael, and then move from there. So, Michael Street, please.

MR. STREET: Okay, thanks. Now, that's a really good question. Almost, well, a big part of what NATO does is focused on interoperability throughout its lifetime. We've had a huge amount of work in creating common standards for everything from, you know, fuel connectors through to digital message transmission formats. So, interoperability is a key aspect of what NATO does. I think of the key benefits that it brings to the nations.

We are involved in some work right now which is where we just first of all doing a little stop-take with nations of what activities that they're currently involved in. What some of their priorities are. But a key part of that is how can we collaborate more effectively? How could we improve the -- interoperability is maybe not the right word that we'd use coming from a data science and an AI development perspective, but essentially, it's how could we improve the interoperability of all of the components that we're using when we're developing AI solutions, and the outputs that we get from AI solutions.

So, for example, it could be around datasets. How could we take datasets and make them readily available in a way that makes sense to people? We know we've lived through bitter experience just to take some data that you've collected and throw it over to somebody else is really, really unhelpful. Especially when as you're developing an AI solution, you will go through a lot of data cleansing, data preparation, maybe data labeling. You'll have a very good picture of what other parameters apply to that dataset when you're using it in a machine learning or an AI development context.

So, how could we make that kind of information more readily understandable by everybody in the alliance so that we can share those big datasets more readily. Maybe so that we can pool those datasets so that we have, you know, instead of 30 disparate datasets or a larger one that gives us that bigger data. It makes the statistical models effective. So, could we make the data more interoperable? Can we make the models more interoperable? We're already involved in some work, actually, with the U.S. and a couple of other nations about sharing models which where we've taken public sector or public domain models and then retraining them and optimizing them to work with military type datasets. Then when we've done that there's a lot of interest from nations of how can we share that? How can we understand what you've done without needing to go through the process again?

And then, finally, when we've developed an AI solution, how do we make sure that that -- or how do we make that interoperate with everything else it needs to fit in with. That might be other AI solutions from other nations. But more often it's actually how do we integrate that into some of our functional tools, our command-and-control tools, our intel tools, in order to make life better and more effective for the end users, for the analysts, the commanders.

So, that interoperability not just within the AI development domain, but also with the outputs of an AI model or an AI tool. How could we integrate that into everything else that we have certainly in the digital domain, physical domain too, so that the end users really see the benefit. Rather than our AI tools being a distinct single isolated little tool that they may go and consult, but then they need to take what they've learned from that off and apply it to the rest of their data we wanted to integrate this into our daily business, you know, our business-as-usual activities.

And that is a challenge. As I say, there are some activities that are ongoing. This question, you know, which will be going out to nations and through the usual roots in NATO in the next -- it might actually have just gone out -- but certainly in the next week or two if it hasn't. Just to get a stop-take to make sure that NATO has a good understanding of what the priority areas are in the nations and how we can best pool that best practice. How we can focus on the bit switch and maybe some of the gaps that get overlooked between different national perspectives. How could we bring that multinational collaboration aspect together. But, absolutely, interoperability is a big part of that. Thank you.

MS. SISSON: Great, thank you. Michael, actually, Michael Gibson, I want to add a little bit of a twist, if I can, to the question for you. Which is, you know, taking those technical requirements, especially given that, you know, different nations and organizations have different capabilities, we can think about all that set of issues that Michael Street has identified for us. When we think about interoperability at the ethics level and the application of principles level, when we talk about AI and ML capabilities being developed, and they can become interoperable that way, what is necessary for that interorganizational trust in the ethical basis of those choices? Is it a matter of transparency that's going to required? Are there going to need to be sort of collaborative, not to trivialize it, but tick lists, things like of that nature? What do you think about when you think about the interoperability of both the tool and its ability to operate, but also the principled foundation that it's built from?

MR. GIBSON: Good questions. So, I'm a policy person, so, I've very much thinking of interoperability in the classic sense of you've got your military on the ground, how do they have the right levels of trust in the unit that's on their flanks, et cetera? I think there is a certain degree of transparency. There's roles, obviously, that NATO have in helping to expose and helping different countries to understand the basis, the policy basis, the assumptions, the ways that certain systems are being configured, et cetera.

But I think it actually goes a little bit deeper than that. There's elements of risk here that we need to understand. I forget who mentioned TEVV. But, you know, what process have they been through? What is acceptable levels of risk for the use of some of these systems? Within the U.K. we have a principle, a law, as law is reasonably practical. You can't eliminate risk completely. But if we don't understand that on the foundational level, and it's really difficult to have that sort of trust in the people in your flanks.

I think there's lots in this about culture. There's lots in this about how our people are trained. And how our people interact. I mean, we do lots of comments. I'd like to see more of them. I think quite frankly we need to be just promoting a seamless interchange in very many ways from the policy layer down to the technical layer in order to just get some deeper understanding. I mean, forums like the AI Partnership for Defense really, really helpful as a show and tell kind of organization. But you do need that deeper level that AI PfD hasn't quite got to thus far.

I think that's the sort of mechanism that's going to help us to, you know, get underneath some of these questions about, well, how is your system trained? What's it actually intended to do and not do? And what data is it being used? What are the assumptions around it so I can actually have a certain level of confidence and trust? What security is around it when you're exchanging data across the boundaries, et cetera? I think fundamentally, it does come down to the people. It comes down to the human systems integration points that were very helpfully made earlier.

And the last thing I'd say is something that the biggest challenge in my eyes, and we talk about our people being our greatest asset, biggest challenge in my eyes to actually operationalizing using these things at scale and at pace is sweat. We really struggle to get the right sort of sweat into our organizations at a technical level. But you need that sweat at every layer of the business in order to, you know, develop and operationalize this stuff.

So, all the way through to, you know, what sweat to you need for insurance to ensure that an AI system is properly developed? What sweat to you need as a leader in order to understand what, you know, the opportunities are and how it can be used and when you're being sold snake oil. What sweat do you need as a user? And then with the interoperability point, a) how do we exchange some of that sweat between us? And but b) how do we really ramp up the exercise in the practicing in practical terms so that we do build up that level of confidence and trust?

MS. SISSON: Great, thank you. Jared, if you had something you wanted to add on any of those matters, please do. I also thought it might be useful if you wanted to offer maybe a case study or an example of the ways of which you've applied these ethical AI guidelines through DIU.

MR. DUNNMON: Sure. Yeah, a couple of comments. I mean, I think, a lot of the pieces on interoperability were on point. I'd say kind of at least one angle that I'd give on it is that there are a couple of different perspectives on it. You know, there's a technical piece of interoperability and then there's an operational piece of interoperability. From a technical standpoint, a lot of that if you think about how that’s done in, you know, for instance in industry and at large scale, a lot of that involves having the right infrastructure.

So, instead of having to, you know, develop a machine learning model and then, you know, kind of have a specific bestow kind of way that we deploy it and the way that we do test evaluation. Having at the very least, standard sets of infrastructure where you're not saying like thou shalt always do it this way, but having a set of tools that folks use that are kind of broadly understood, broadly used, and kind of reasonably in alignment with best practice, is really helpful. This is kind of part of the whole, you know, ML ops argument to say, look, I have a model and now there's all this kind of care and feeding I have to do for this constantly that involves, you know, TEVV, that involves post-deployment monitoring, that in having standard, you know, and again, I don't say standard with a big S, I just mean having a commonly understood way and process by which we evaluate models. By which we test, poke, prod, test, and evaluate them. By which we then document -- and this is getting to the second piece -- document the risks associated alongside kind of specific error modes that we find, and then go and say, all right, here is the risk/reward profile that we're seeing here, operationally, do we think we should deploy this?

And this is where we get into the, you know, away from the technical pieces, which are important to be able to communicate and say, yes, you know, your definition of an apple is the same as my definition of an apple. But then there's this operational angle under interoperability. Which really has to go to all the points on human computer interactions, the points, you know, on training. But it's also on documentation. I mean, and a lot of this is a matter of saying, all right, if we are doing a rigorous job of when we do test evaluation, verification, and validation, documenting exactly what our assumptions were, exactly what we found. Being clear about how we're measuring those things. And then on the back end saying, look, here's the performance envelope we're seeing. Here's the profile we're seeing. You know, I'm going to make something up, but sure, maybe this will work 90 percent of the time. And we know it's going to fail 10 percent of the time. Like we are aware of that.

Now, in certain operational contexts, that may well be a worthwhile -- that may be a viable risk profile to deploy something. But then if in practice when we hit, you know, the 10 percent and it fails, we don't look back and say, oh, well, gosh, our entire system is broken. It's like, no, we thought about that. One of two things happens. Either a system fails. We look back and we say, no, we thought about that. That was part of the decision to deploy this. And, you know, we either need to adjust our calculation of risk or adjust our assessment, say how often that failure is going to happen or how severe it's going to be. Or we look back and we say, no, actually, we look back and say, we didn't think about failure mode. Okay, we need to really rethink this.

Or we look back and say, look, this was part of the process, you know, from a statistical standpoint. We knew this might happen. We thought about it. And our process was fine and we actually may even consider continuing to deploy that because, again, the alternative may be substantially worse. So, that's where this operational part of interoperability and making sure that the documentation is clear, which has a tieback to making sure that we're using the same tools that define kind of apples in the same way, is really important.

And I can give on the case study standpoint, I can give a, you know, a really direct example of that. And this actually, I think is a really interesting one. One of the programs that we actually describe in the white paper is a program where we're working on leveraging kind of all sorts of different types of publicly available, commercially available information to build knowledge graphs that allow folks to better track transactional criminal organizations.

So, one of the things that you when you do that, you're basically building machine learning models that go through, you know, instead of having human read millions of documents and try to figure out, you know, John Smith over here is John Smith over there and those two are the same person, so, this piece of information is relative to this piece of information. What you do is you train, you know, machine learning models to go and identify entities. So, name people, you know, names, you know, people, places, and things. And then you build many different relation extraction models.

So, if I have two entities in a sentence, you know, Bob and Jane, you know, you might ask in these model, this relation extraction model, based on this sentence, are Bob and Jane married? And or is Bob an investor in Jane's company? There's all sorts of relations you might care about. And what you do is you then build up this graph where if you look at a node, a node might be Bob or Jane. And, you know, you can find, oh, well, this is Bob. Bob is related to, you know, Bill. Bob is invested in Joe's company over here. You can very quickly start to understand the world in a much more digestible way for a human being than you would, you know, then to just kind of read your documents and hoping you found the right thing.

So, if we think about that and walk through some of our art process, you find some interesting things. So, first, you know, what are our metrics or tasks in our baseline? Well, there's actually a cascade here. Because at the end of the day, your operational task is very much, for instance, can I identify behavior or kind of sets of relations that might tip me off to criminal behavior? That's one thing. And that's a human process where you're kind of looking through, you're trying to surface the information that's relevant. And so, you know, your metric might be, you know, out of a, you know, known set, you know, out of a known set of, you know, set of instances that you want to make sure you find, you know, how many of them are your humans actually able to finding using the system? And how fast are they able to do it?

That's a high-level metric. That's like an operational metric. Now, you have to dig in. But underneath that operational metric, there are probably hundreds of different machine learning model structures. And so, how are you measuring performance of each one of those and how does potential failure on each one of those actually effect the downstream task? It's a tricky question to answer actually. And so, that's why it's really important to have these continuous monitoring procedures in place to make sure that you're continuously assessing that.

You know, there's this question again, here, you know, ownership, access, provenance to data, right? You know, where is the data coming from? Are we confident that our sources are as accurate as we thought they were when we started? These are things that we have to continuously assess. You know, who are the end users here? I mean, who are the stakeholders? You know, end users are folks who, you know, obviously the analysts. But they're also, you know, there's a bunch of stakeholders here, right? Anyone who could be affected, you know, if there's an investigation that ends up being run. Like anyone who could be affected by that investigation is a stakeholder in that algorithm. And so, you've got to think about those, you know, those potential costs, all of those things.

You know, rollback and air identification. You know, these are models that are being developed, you know, if they're being developed by a private company, and you have to be constantly retraining, you know, have you, you know, how do you redeploy something? You find a model and you decide today, well, actually, we find this thing's not working right now. How do we roll it back? And if you're using commercial tooling, it's actually often not that hard. You kind of revert. That's a standard practice. Sometimes in government software we don't design things where it's that easy to do.

You know, but then identifying errors. You know, is there a person or a auditor that is consistently whose part of their job is to make sure that they're looking at the outputs that are coming out of this algorithm or out of this knowledge graph and saying, all right, you know, how are we continuing to assess quality here? And by continuously running tests and evaluations.

And so, these are things that we talk about concretely in the process of that program. And once it's deployed, again, you know, this is one that we haven't deployed yet. But once we get to deployment, we're going to need to be able to scale those processes. You know, and one argument would be look, you know, we've seen programs where this happens, you get to a point where, yes, the things works. However, you know, the care and feeding of the model has not been sufficiently resourced. Or the cost of doing so is actually so high, that it may actually not be worth deploying this thing.

And to be clear, and something just to take away, like that is and should be a viable answer out of all of these types of -- they shouldn't be just like a -- I said a tick box where it's like, yeah, yeah, we did that. Okay, now we can go and deploy it. I mean, part of the purpose of these, you know, of the case studies as well, and it doesn't apply to this particular case, but one of the outcomes here should be able to say, look, we found an application that we think had promise. We got to a certain place in development and found out, yeah, actually, either it's going to cost too much or the risk profile isn't what we accept. And we're going to debug that and not do it. And we're going to do it a different way. And that's actually a very positive outcome. Like deciding not to do a thing that's not great is a good idea.

So, part of the process here is to make sure that we have those conversations early as possible. So, if you go back and say to someone, okay, well, what's your metric for success here and they can't answer that question, I mean, you're not going to really develop an efficient system. If you go back and say, all right, well, where's your data coming from and they're like, oh, actually, I kind of think I know where it's coming from, but we're not really sure. And, you know, we'd have to really go and dig in to figure that out. Again, that maybe be a cost that is so costly that folks decide to do this a different way or assign human analysts to do something.

So, this is all just to say that I guess this knowledge graph construction case is really interesting because you think about it, it is an end-to-end task. And in terms of trust, like again, folks were talking trust here, one of the ways that you can do that technically is try to surface, you know, when there's a relation in that graph, for instance, you not only say, hey, there's a relation in this graph, you know, Bob is invested in Jane's company. You actually surfaced the document that said that. So, a human can look and say, okay, do I actually belief that. Without those types of kind of verifiability, it's really difficult to deploy a system like that because it's built on, you know, hundreds of models running constantly.

And so, the thing to communicate here as I wrap up is that, yes, we talk about this a lot of times in the context of, you know, a single model, like there's a computer vision model that does something or, you know, a machine learning model that does something. In reality, a lot of these systems are going to be massive numbers of interacting computational kind of objects and thinking about not only how do we do this once, but how can we do this when we have a situation where we have hundreds of models running is a really important problem that kind of we need to think about as we start to look at scaling these applications within, you know, the defense apparatus. So, I'll stop there, but thanks for the question.

MS. SISSON: Great, thanks very much. So, I think pretty much any observer, casual or otherwise, it would be understandable for them to react to all of the content that you've all raised and say, wow, this is really hard, right? There's so much complexity involved and challenge and difficulty. And yet, all of you are here because, clearly, you're optimists about this and you think it's important enough and feasible enough to keep working at it.

So, I want to ask a question about the challenges that you just raised, Jared, and including going back to comments about documentation and I'll start by bringing in Heather. What I want to ask about the number of issues and that you've all raised, the work that has to go into creating these kinds of systems in a way that does engender trust in the direct users, but also the societies that support and sustain them. How do you think about time and urgency? And reconciling the detailed level of things like documentation, for example, that has to go into putting ethical principles into practice but also the pressures, external or internal, based on the rate of change in technology, the need of policy makers, and the militaries that are going to use these tools and systems as they emerge, how do you reconcile those two? Do we make too much of the time pressure? Do we make too little of the time pressure? How do you think about striking that balance?

MS. ROFF: Yeah, Melanie, thanks. Just two quick points. I think on the, I mean, obviously, I think documentation's a pretty big deal. That's why I put it in the whole, you know, traceable principle. It is good systems engineering, right? But I also think that one of the things about saying documentation thing, which I think is completely necessary, we have to also kind of continually stress within the DOD and other militaries is that people don't read documentation all that often. And so, getting back to our like human system side of the house of really thinking about what our people will do.

It's one thing to document it and then have a failure or in DOD parlance, a mishap, and then catalog the mishap. And assuming we can catalog the mishap in a correct way to get down to the bottom of it. But the documentation, I think, we really need to not just document it, but widely disseminate our findings and lessons learned and things like that.

Which we are really not doing a great job of widely sharing that information so that we aren't reinventing the wheel. And then putting it into -- we can network that a lot better than we do. And we need to do a better data infrastructure. Our data modernization strategy lays some things out, but we still suffer with like 12,000 data lakes in the DOD alone. So, there's a lot of different problems there. So, I think that documentation rather than just creating more noise, needs to be streamlined in a much more coherent fashion. And then that can be automated as well in different ways so that we can actually bring to attention where there's so much overwhelming information.

Then in terms of thinking about how we can learn lessons going forward, you know, one of the things that I think is really, really important is to continually push on where the incentive structures lie. So, if we think about complex systems failures, which is what we're going to be talking about now, right? We can go back to a case, a very familiar case of the 1971 Ford Pinto that was rushed into production, reused a chassis that was inappropriately used. And after doing multiple risk assessments and investigations and accidents after people were burned to death, found that, you know, this bolt here and this bolt here, well, we reused it because it was a cost factor and there was a speed factor and there was a gas shortage factor and there was this complex of things happening. They ultimately found that even the payout to those people was going to be less costly than a recall right?

So, you had the incentives in these kind of misaligned ways. And so, I think it's really important to remember that even our metric of success we need to be from an ethics perspective, pretty clear about what our metric of success is here and where we're going forward. Because if we just say, look, this is too expensive, yes, it might be very expensive, but we also have to think about what is our metric for morals, right? So, it might not just be lives saved. It might be nuclear stability. It might be economic prosperity. It might be human welfare. All of these things. But we have to be quite explicit about what those metrics are and then be very, very explicit and open to the way in which we're bringing our assumptions to those benchmarks and those metrics and how we're operationalizing it. Because otherwise, we might end up in the kind of Ford Pinto case of, you know what, it's way too costly to recall it. A few more dead bodies is okay. So, I think that that's something that we really need to keep our eye on the ball on. Thanks a lot.

MS. SISSON: Okay. Thanks, Heather. And my apologies to the panel and the audience. My eyes were bigger than my mouth here in terms of time. We are at our appointed hour. So, before we close, I thought I would express the thanks of the Brookings Institution and to our audience that I know has learned a lot from you all and I thought I would maybe do a real quick lightening round if any of you has any last comment that you can't hold back and we'll sign off from there. So, maybe just back from the top, Michael Gibson, if you had any last words that you wanted to add.

MR. GIBSON: No, nothing in particular. I found that last comment about what's our metric for morals really quite interesting for how we go about operationalizing these ethical principles and actually making them more than just words on the page. Making things that people actually breathe and own. So, no, that's one of my takeaways, thank you.

MS. SISSON: Great, thank you. Michael Street, thank you.

MR. STREET: Thanks also. A couple of things that we've learned through working through this is collaboration. And particularly multinational collaboration bringing different, slightly different viewpoints and perspectives is really helpful and really important. And, yeah, just as Michael Gibson had mentioned, words on a page. One of the things that we find when we're dealing with natural language understanding is the ways things can be interpreted. Words on a page can be interpreted in different ways. So, ways that we can document or at least describe what things do, some of the boundary conditions, things like that, is really important. But if we can do it in ways that are clearer and better prescribed than maybe just a written document alone, that's also something which could be useful, a) for validating that systems perform the way that they're meant to perform, and also, for analyzing all the data and the documentation that we go to generate about those systems because at some point we've got to want to do that too. Thank you.

MS. SISSON: Great, thank you. Jared.

MR. DUNNMON: Yeah, I'd just reemphasize the point that I think was just made about, you know, how do we make these processes things that are breathable and that folks can have constantly in the back of their mind and that requires, you know, disseminating those findings, as was just mentioned. And thinking about the best way to do that, the most effective way to do that, and the most efficient way to do that so that folks aren't, you know, getting drowned in words and kind of look at some of these ideas and say, ah, you know, this is some giant, you know, packet of documentation that I don't know what to do with. That is a really important point. I think it's kind of the exciting evolution of the next things we need to think about here. So, you know, thanks for all of the conversation here. I really appreciate it.

MS. SISSON: Great, thank you very much, Jared. Heather, last note and we'll go from there.

MS. ROFF: Yeah, I would just say on the last note is that ethics is a continual process, right? It's never complete. It's an approximation. It's an asymptotic line. So, whatever you're doing and however you're doing it within your perspective place or organization is it's a continual act of critical thinking and that's how you should appreciate it and approach it. And don't think about it as box checking.

MS. SISSON: That is, indeed, an excellent last note. Thank you all again on behalf of the Brookings Institution for joining us. And to all of you that tuned in, we greatly appreciate it. And hope everybody has a wonderful day. Thank you.

* * * * *

CERTIFICATE OF NOTARY PUBLIC

I, Carleton J. Anderson, III do hereby certify that the forgoing electronic file when originally transmitted was reduced to text at my direction; that said transcript is a true record of the proceedings therein referenced; that I am neither counsel for, related to, nor employed by any of the parties to the action in which these proceedings were taken; and, furthermore, that I am neither a relative or employee of any attorney or counsel employed by the parties hereto, nor financially or otherwise interested in the outcome of this action.

Carleton J. Anderson, III

(Signature and Seal on File)

Notary Public in and for the Commonwealth of Virginia

Commission No. 351998

Expires: November 30, 2024