Bryan Clark  0:00

CDAO is sort of central, like Aditi was just talking about to a lot of these integration efforts to bring new capabilities to bear to support warfighting needs of today. And, you know, Margaret Palmieri is at the center of a lot of those efforts as the deputy leading several initiatives that we'll talk about today. But thank you very much for being here. We appreciate it. So, you know, we just talked with a DD about, you know, how we're trying to bring a lot of new players into the defense ecosystem and try to align solutions with things that, you know, the combatant commanders or commanders in general, are looking to address. Yeah. What's the what is the CDAO's role in that process? You know, generally, and then we can talk more specifically about some of the initiatives you guys are pursuing.

Margaret Palmieri  0:45

CDAO, a relatively new organization, at least, it doesn't feel that way. But it is. Step in June of 2022. Our mission is to accelerate the department's adoption of data analytics and AI, so not that big of a mission. But, you know, our focus is on that question of acceleration of adoption. And we know, inside of DOD, one of the largest organizations in the world, 3.5 million people, over 160 Different countries where we have a presence not to mention under the sea and in the outer space. And the Pentagon is not going to be the source of all the innovation that could happen in that type of ecosystem and environment. And so accelerating adoption really means making government data accessible to both government and industry and our allies and partners in order to create capabilities necessary to solve DoD problems. And the more we can do that, across multiple levels, and across multiple partners, the better and so we have a wide range of theories and hypotheses of how to make that happen. And we take a really experimentation based approach, to learn and to codify the things that work.

Bryan Clark  1:54

So obviously, this requires, you know, some pipeline basically, of good ideas, you know, the operational concepts that we want to empower, and then maybe some material solutions that are part of that. And then, you know, a process by which we sort of assess how they can come together in interesting ways, and how we can sort of layer on top of that artificial intelligence and data analytics and interoperability. So what do you what's the engine that you guys are working, are developing to do that.

Margaret Palmieri  2:27
So I'll talk about two things. One is a little bit more abstract. We have a hierarchy of needs, as we call it, we put out a new DoD data analytics and AI adoption strategy just a couple of months ago, in there, you can read all about the hierarchy of needs. But basically, it's not like a Maslow's hierarchy where one is a necessary precondition for another, but it helps us think through the maturity of the department or the maturity of a project, when it comes to delivering capability. So at the core foundation is data quality, lots of people love to jump to the algorithm first. But really getting getting access to that data, making sure you understand it, you understand where it's not high quality, where it is high quality, what it's good for, what it's not good for, is it labeled and tagged. And all that fun stuff for the project that you need to do is step one, and unfortunately, I'm probably your most boring guest, inside of this conference when it comes to that. Because it is just really specific work the people who like it, love it, which is so wonderful. And we've got a ton of them in the department that are just really enthusiastic, because when you get that right, you unlock the two top, the top two pieces of the pyramid one is analytics. We like to put metrics with that, because a lot of times in order to really get our result in a project, you have to understand how DOD does business. And metrics are a really great conversation, not like the, you know, typical metrics that we get, which are several pages of very discrete things, but but really an understanding of how does the system or how does the process or how does the workflow work? And how do you know if you're actually improving it. So if we want a different outcome, we have to have a sense of how we do business today. And a lot of that in DOD is passed on person to person, it's like knowledge you acquire by being around for a while, which, you know, is good in some extent, but not necessarily replicable or sustainable. And in other cases, and, and so we look at analytics and metrics kind of in that block. Sometimes people ask for AI, they just need a dashboard. Or they just need a simple like regression have these valuables relates or a driver tree. And then at the top is AI and once you have a sense of you know where you want to play there. You can put put AI on it as well. Around the pyramid, we have this thing called foundational enablers. And so these are all the processes, talent management, acquisition, policy type things that have to happen to make this repeatable across the department at scale. So that's kind of how we do that more broadly. And then we have specific projects that we take on from on everything from business metrics to Gen Z to and. And that's how we kind of think through how we're going to approach those different efforts.

Bryan Clark  5:10

So we think about that pyramid. You started with the data, you know, the data quality. Bar daughters as data sciences, she loves. She does. She loves doing that. That's right, she will connect you. Right. Yeah. So, but she, but so that's obviously something that's absolutely essential to anything else. And like you said, so? Are there ways to automate that process to some degree? Or does it have to be me? Because it seems like it's really a slog, and a manual process? In a lot of ways? Yeah,
Margaret Palmieri  5:41

I think there potentially could be, but the challenge is, data scientists have this incredible skill set as translators between between mission and data. And if you find the good ones, they not only just you work through the, you know, the slog of tagging data and or labeling the right columns and things like that. But but they actually translate what fields are really valuable to the operators? Or where should we focus our efforts to improve high quality data versus other fields that we might not care about? There's, of course, I think, over 4000 data systems and said, the Department of Defense. And so you can imagine that data appears the same in different ways and multiple systems. And so So how do you actually rationalize that and correlate that together? And so without a doubt, you can write scripts to do some of that, once you understand that, but I think, at least, you know, actually, probably forever, we will need the the thought process of folks that can do that mission to data translation. So

Bryan Clark  6:36

to what degree do you have to have that as an interaction with operators? I mean, does this end up having to be something that you have to embed like we were just talking about with a DD embedding people with those operational units to say, Okay, let me help you, let's translate your data into something useful?

Margaret Palmieri  6:51

Yeah. Ideally, there's there's two approaches that I think have to go together. We do this through the AI data acceleration Initiative, or ADA, where we put people at combatant commands with data science and algorithmic expertise. But we also have a hub back at CDAO. And so why don't we do that? The department is very large. We don't have to solve the same problem multiple times if we can avoid it. And so yes, I think you do have to put the experts. Without a doubt, I don't think I know you have to put the experts inside the different organizations, they have to understand the workflows. Even if you can create the same algorithm across the same problem set how people actually ingest that into their workflow is really important. And what we see we do a bunch of experimentation, we get some really good results and experimentation. But if we don't see those same workflows, replicated in actual operations, you don't actually get the same out value. And so you get the same application, the same algorithm. It's available, but it's not used to the full extent. And so we keep a hub back at CDA to, to look at data quality across the department. Our director Craig Martell has a concept from industry called called core concept data, which is cocoa, but it's not contractor operated anyway. But it's this idea that 20% of your data unlocks 80% of your use cases. And we generally find that right? It's personnel, it's
logistics or readiness information. It's your people, your stuff, your locations, they they tend to be similar across the board. But so we keep that at CDIO. So that we can look functionally at aligning communities across different organizations. But then we have to go to the place where the work is being done to make sure we understand it.

Bryan Clark  8:31

So to what degree do you find that an organization or operational command has to change its workflow to make the best use of the data and the way it's structured? Today, because I feel like a lot of this and, you know, we started out as a one way conversation of the warfighters got a certain way of doing business, engineer this data. So it's important to the way I do business today, when in fact, maybe it would be better if they just adapted their workflow to a much easier way to use the day.

Margaret Palmieri  8:56

Yeah, it's a constant give and take, I would say, I'm trying to guess the percentage here to be data driven. But you know, at least 90% of the time, there's benefit from changing the workflow. And that's not always clear upfront, I'd say 10% of the time, that's clear upfront. And so in this process of bringing new technologies, new data and analytics into a current workflow, you may start off with, Hey, let's, let's help digitize some of what you do. We're in the process, most people realize they can get so much more out of it if they started to adjust. This is so hard, right? Because in an individual workflow, and a lot of times we pick the user use case where you know, you've got one user working through something, it's easy to digitize that automate that when you start bringing in the handoff of information, which is really what helps enrich our decision making is that we get diverse perspectives on some information, and then we're able to bring it together and present options to decision makers. Well, we have different perspectives on different data. And so as those workflows start to integrate, not only do You get the normal process, disagreement, but you may also get well I like the way we do it now Well, I like this way why I like a different way, which is why we we experiment a lot instead of video.

Bryan Clark  10:09

And then when we go to like a jazzy to type of idea, like we're trying to operationalize this with a group of units. So then it's some of it is not just kind of the predilection of somebody to operate in a certain way. But it's also some of this is hardwired into these machines, or at least programmed into the machines in a way that's not easily changed. So how do you navigate that back and forth?
Margaret Palmieri  10:32

When one step at a time, I think it really is, we take a very mission lens focus, and we'll, we'll start off with specific workflows, one of which we have, of course, coined a new acronym that has many, many letters, so many letters, we put a number in it. So combatant command and combatant command and coalition. Oh shoot sci fi, sci fi, ve its combat and command coalition collaboration and planning. Okay, see five feet, there we go. What is that, we know from the global integration mission that has been given to the Joint Staff for a couple of years now, that pulling together decision making across combatant commands, especially on you know, whether it's global, or whether it's multi regional, is really critical to informing the chairman and the Secretary's decisions, and advice to the President. And as we look across those workflows, there are you know, really Microsoft Office, which is wonderful based processes for doing that. But to digitize those workflows, creates so many different opportunities, because awareness is extended across multiple organizations, the type of options, the number of options and creativity of options is changing as we bring those groups closer together. And, and then we're really finding out, you know, which applications are ready to jump in and which, you know, are digitally ready, which ones are not digitally ready. The cool thing is, in some cases, you can just simply go in and change how metadata is exchanged and received, and you get a completely new capability that you don't have to go through any sort of J rock or palm process. Yeah. In other cases, I think as we get closer to the tactical edge, the physics and the weapon systems are going to be a little bit difficult, right? Yeah.

Bryan Clark  12:20

Right, right. So when you look at the, when we look at the idea of one, there's 1000s of data systems in the department, so how do you triage sort of, well, what am I even going to tackle because some things are probably going to be, you know, systems that hopefully we will eventually walk away from, like a fitteds are saying things, although we've said that for decades now, but how do you sort of triage which data systems I want to pursue and which kill chains or which workflows I'm going to tackle.

Margaret Palmieri  12:49

So we focus on the global integration mission threat, and we focused on long range maritime fires, specifically for the Indo Pacific as our first two mission threads, which are joint and combined, and so encompass a wide variety of systems, even though they they sound like they're narrow. And for DOD they are, they're discreet enough where we can actually and we've done two global information dominance experiments or guide events on this, we do a guide event
every 90 days. In guide seven, we actually just digitize the targeting data through the workflow for that second mission thread from Can we go from the combatant command all the way down to the tactical level? Can the, you know, shooter actually Roger up that they have it? Can they fire weapon off the quality of data that they've received? That was wonderful just to track the data paths from again, like the geeky perspective, but now we know what works, what doesn't? Where do we have to go after in guide eight, we were able to do the same thing with even more weapons associated with that kill chain. And we're able to look across the joint force at you know, where different systems needed to make adjustments. And that's how we do it. And we give the feedback to the programs and we take a look at, you know, is this a simple change? Is this more, you know, hardware or FP related thing?

Bryan Clark 14:05

So, when you do that do, is there an option, then? I mean, once you've digitized that workflow, can you turn around then and look at ways to virtually assess it to see how can I? What are the things that seems really make a difference here and which things don't and to guide sort of program decisions? Like if I want to go build new stuff, or add new stuff? What's the most effective things to add? So let me kind of virtualizing that machine, if you will, that like

Margaret Palmieri 14:26

like in a modeling and seminar? Yeah. I'm sure you could, and I'm sure the services have in fact, I know they have in many different places. How we bring that together as a as a joint force, and especially at the operational level and above? You know, I think we we get a lot of that through experimentation, but absolutely more through virtualization. From that perspective. Yeah.

Bryan Clark 14:48

So I'm gonna just, uh, let the audience know we're going to open up to questions here in a couple minutes. We'll we'll go to Fargo What time do we need to start the panel to so I was 1415. Right? Okay, two minutes. Okay, got it. All right, I just want to make sure I kind of where I am. So I will open up to questions here in a minute. But, um, so you you talk a little bit about international partners. So, you know, we're we do projects with the Japanese, the Australians in particular right now, how are you? How do you how do you work with them to look at ways to sort of make your digital fabrics work together? Because right now I get the impression that there's unless they buy the exact systems the US is buying? There's not really that easy interoperability.

Margaret Palmieri 15:34
Yeah. So we've been focused on Five Eyes mostly, first, the UK has been a partner in guide, I think, since guide five, excuse me and guide eight, we actually were able to successfully pass targets over across native systems, and show what we can do just by having it open interfaces. And the same definition for you know, what's the target? What What information do you really need? And what does that look like? That's a that's a small step to where we really need to go, I think in terms of digitizing and being able to share all the information that we need to I know Australia has been an observer and a couple of those exercises. And what we're trying to build inside of Cabo is a we'll call it a spec for for how to exchange data, we really want multiple companies to be able to come in, talk with a combatant commander, talk with a user in the field, get a sense of a problem and very, very rapidly be able to get a contract and get access to government data and access to government networks and be able to solve it in, let's say, days, maybe maybe weeks. But definitely not. If it's really hard, maybe months but but definitely not the years it takes at this point right now. And so we have an acquisition arm instead of CDA, that's looking at reducing barriers and working with diu. On on entry there, we got a really creative authorizing official from the Air Force that believes in reciprocity and risk management and not checklists to really help us think through, you know, how do we keep our data secure, but also make it available. And then this last piece, on the technical side is this data integration layer, it's not really layer. It's a collection of data mesh services, it's a collection of data accesses, where we can reduce the barrier for anyone who has a problem to be able to find the data that they need, and apply it to their problem set. And make sure that once they bring it into an environment, whether it's industry provided or government, their the enriched data product can be shared back out. And that's really what we care about, we want to make sure that the government, we often talk about government ownership of data, the ownership word is really about the accessibility and our operability across the force, and that it's not, so piped or pushed into a corner somewhere. But we fully want to leverage all of the the richness that industry brings to make products better.

Bryan Clark  18:00

So when you think about how you acquire these capabilities, or work with software providers, for example. So I think which of these is I would interpret is the government is owning sort of the interfaces between these different systems. And then you may have multiple providers that are actually delivering the software that works on the opposite sides of those interfaces. So So in a lot of ways, are you still letting them retain the kind of ownership of that software in a way that allows them to basically, you know, have their intellectual property and be able to have a business model that relies on it? Yes,

Margaret Palmieri  18:33
absolutely. We want companies to make money. We want we want people to come to America with their innovative ideas, start a company make money, and enjoy working for the government. And we're working on all those things. No. Yeah, without a doubt. What we don't want is just the raw data back out, we want the enriched data. And so I think where the difference has been in the past has been, you're gonna have your data back, but anything that we did to make it better inside the ecosystem you can't have and so we're saying no, no, we want we want all that back. Ideally, you know, American innovative commercial talent is one of our hugest assets, a strategic assets inside of the US. And the more we can support that ecosystem and cultivate that, the more we want, and ultimately, you know, we can get interoperability through how we define what we need and enter those interfaces. Yeah.

Bryan Clark  19:32

And so, yeah, that sounds like so for my for if you're a software company, that sounds like a great model to be able to, you know, provide services and capability to the government. And still, you know, not and not have to get your stuff absorbed into some larger lead system integrators contest system, for example, or whatever. How do you Yeah, how do you make it so that you All these individual companies can be managed separately do when you when you have your new authorizer for contracts for software are you having to make sure that you've got basically vendors that are providing a kind of ecosystem sort of service are different than those that might be providing the control the AI enabled control system for particular unmanned vehicle?

Margaret Palmieri  20:21

Yeah, you know, we talk a lot in CDA about the difference between technical decisions and business decisions. And so, the technical underpinning, I think, is pretty straightforward about how to design something well, or to be interoperable. And an accessible, the business decisions are a little bit more complex, because we do need companies that can provide sustainability and reliability and helpdesk support. And, you know, they might not be the same companies that are gonna give you the next most creative algorithm, right. And I think that's that, but that's something that we also want to experiment through. We want to work with the acquisition arms of the department to really figure out best practices there. And I think you're gonna see a bunch of different models throughout DoD as we learn through that. But we're not ready to kind of say, hey, there's a centralized way to do it. But just be really thoughtful about how you combine innovation and sustainability and reliability. Cuz our mission is warfighting.

Bryan Clark  21:16
Margaret Palmieri  22:00

So we have said that the work that we will do will really stop at that operational levels to tie the combatant commands with the Joint Task Force commanders. In some cases, those joint task force commanders are service components. But we worked very closely with Project overmatch project convergence, ABMS, and C three, BM office and Air Force to make sure that the work that we're doing will enable them to also use data that's coming from different parts of the theater. But we've left most of the joint Kill Chain development to the services to really work on and we think that's a good division of, you know, the the business piece, the services have a lot of those pieces as well as expertise. But in guide, we have fielded a couple applications just to be able to test whether or not the data is flowing, because we need to see where it goes when people have to interact with it. And so system Yeah, so we've got a couple of the experimentation realm that are exciting and promising that, you know, the services can definitely look at and

Bryan Clark  23:02

learn probably, but it's about getting them the data, because fundamentally, if they don't have that ability to share data across multiple domains, system services, you're not going to be able to make you're not opening up those kill chains or those courses of action. So let's, let's take some questions from the audience. So a few of you, ma'am. Yes. So bring the microphone is right over there now. So, state your name and affiliation. Hi, I'm

Speaker 1  23:27

Brandy Benson from defense scoop. Thanks for doing this. Good to see you, Margaret Palmieri. I want to stick on guide for a second. I remember whenever guide eight kicked off big priority for the CDA to was reaching a minimum viable capability for Jad C two. So I'm wondering was that reached in guide eight? And can you talk to us about where guide nine is going? And what the status if that is? Yeah, sure.
So we're going to do an outbreak of guide eight, internal to the building in a week or so. So I'll keep my comments a little bit more vague than if we had already briefed them out. But we did. Now Minimum Viable capability is largely a connection of existing capabilities that we have today that have shared data in new ways, and have brought together a combination of new applications, new data services, with users to create better workflows. Now, we have done that in the experimentation environment, there are a couple of things that we need to consider as we bring them into operations. One is a full appropriation and FY 24, which is absolutely critical. We doubled or more our budget in Cao from 23 to 24, and don't have any access to that funding right now to make that that minimum viable capability truly accessible and robust across the fit up and so we're really excited about the potential for an appropriation the other There is back to our conversation about workflows, the adoption on the the warfighting side and the cop side to now say okay, what we did in experimentation we can do in the real world every day. The cool thing about guide is once you make those connections, inside of experimentation we are we are dealing with live data. We are on live networks, there's nothing experimental or modeling about what we're doing. And so they are available for warfighters. Today, there's no problem of record or POM process change that needs to happen, no additional investment other than the full appropriation and 24. So absolutely, we have a new connection, a new set of connections across multiple data fabrics and applications. In guide nine, we are going to align with a couple exercises the Army's doing in Project convergence for because we really want to see how the combat commands and the Joint Task Forces now take that down to a tactical level with a service, you know, army's leading it but it's it's been a joint exercise for a while, or I should say experiment for a while, army's really creative about how to look at things that work and things that don't work. So they make a really great partner. They're also really interested in the institutional pieces of this. And we've had a couple of conversations from their learnings at a tactical level back to you know, how do we think about some of those systems changes. So what we'll align with them in the March timeframe. And then for guide 10, we will align within a Pickens Valley and shield effort, which I know is also a focus of other capability development, and really excited to align with the global information or on global integration, experiment G that the Joint Staff runs. And so we're really trying to take guide to align with existing activities and start to bring those those data connections closer to stuff that people are already doing.

And that's a great point that we did, I shouldn't probably hammer that earlier. But it's in software is naturally scalable. Once you've created this set of workflows that are, you know, have interoperable data and you've completed an effects chain, basically, it's automatically scaled, and it's a leave behind it. You know, whatever. Yeah. So one more question before, we have to shift
over because by going to the panel, so So one more question. Anybody from the audience have a question? Okay, well, I have a actually, last question. So the, so we didn't really talk about but the other public is getting this crash course in AI thanks to chat, GBT. And you know, all these other Geminis and these other systems, I think it's better generalize better. But in any event, so you're obviously trying to provide some AI enabled tools to the department, mostly, I guess, focused around decision making and thought and thinking about, you know, what are the options available to candidates? But obviously, I mean, you're working the Navy digital office, you know, previous to this. Also, there's a lot of places AI can get used in the department beyond just decision support for leaders. So what are some examples where you guys have been able to kind of push that out to the warfighter.

Margaret Palmieri  27:59

A big one is actually autonomy. And so the joint AI center had started a really interesting project called Smart Sensor, they wanted it to fully autonomous and Q nine, not just the aircraft vehicle, but the sensors. And so for several years, DOD has been looking at how do you really do unmanned autonomy enabled by AI. And so replicators actually just had a really great time, we're looking at, you know, how to do a data hub as part of AI data hub as part of replicator to to accelerate that learning across multiple companies and different services. And so I think the autonomy use cases is front and center in our mind. And to bring those those best practices, we are really focused on the scaffolding we call it which are kind of not necessarily all of the potential AI projects that we could do inside of CDO. But really, how do we enable all of the other organizations inside the government to have access to labeled data and not have to go find that themselves? We are very focused on responsible AI. So how do we do tests and evaluation? Responsible AI, you know, isn't just is the algorithm performing? As we expected, but is the operator aware of when that performance changes because we often forget that once you feel the AI, you have to continually monitor how it is performing and and make adjustments as you go. And the operator is right directly in the center of that. And so the testing, evaluation and how we bring either into contracts or into how the government does business, the continuous monitoring and adjustment and training of those pieces there. And then we've got Taskforce, Lima, which is our generative AI Task Force. I think we're over. I'm gonna get the number wrong but we've got 100 use cases or so have large language module across the department. We'll have our CDAO to annual symposium that's coming up into the end of February, huge LLM track there. And not just in terms of speaking events, but interactions with industry around you. How robust are the models? How do we give confidence You know, to our users, based off of what their what the models are providing back, and we look at that through a different another kind of hierarchy. But, you know, it's different to support HR processes, then supporting commanders decision making. And so our tolerance levels there, maybe it'd be slightly off, but we're thinking through how to enable the department to do those those as well. You
Bryan Clark  30:21

I can think about there's there's 1000s of use cases for employing large language models on the trove of data that the department has available. And so test evaluation, I guess, you know, the kind of the big takeaway there is you have to get away from testing against specific, very well defined metrics and look instead of kind of behaviors and how a system performs, yeah, consistent with how we thought it would perform,

Margaret Palmieri  30:40

You can definitely have metrics. But you know, AI is very environment dependent. We saw this in Project Maven, and in some cases, you know, in the ground war, where, you know, you may come in with certain expectations for how an algorithm performs, because of the data was trained on and then you go into a fight, and things don't look the way they used to. And so now the performance expectation is different, doesn't mean they have no utility, but you just have to retrain and and so there's, there's gonna be an interesting learning curve as DoD goes from hardware to software and now from software to AI. But I think we're postured to do that. Yeah.

Bryan Clark  31:19

Well, I think we'll end it there. So thank you very much, Margaret Palmieri for being here today. And so we're going to take a quick break and reset the stage