SEI Webcast

AI Engineering: The National Initiative for Human-Centered, Robust and Secure, and Scalable AI

by Rachel Dzombak, Matt Gaston, Frank Redner

Shane McGraw: Hello and welcome to today's SEI webcast, "AI Engineering, the National Initiative for Human-Centered Robust, Secure and Scalable AI." My name is Shane McGraw, Outreach Team Lead here at the Software Engineering Institute and I'd like to thank you for attending. We want to make today's discussion as interactive as possible so we will address questions throughout today's talk and you can submit those questions in the YouTube chat area and we will get to as many as we can. Our featured speakers today are Dr. Rachel Dzombak, Dr. Matt Gaston and Mr. Frank Redner. Rachel leads Digital Transformation at the SEI's Emerging Technology Center. In her role she works with organizations to realize the capability of artificial intelligence or mission outcomes. Prior to joining CMU, Dr. Dzombak was an Innovation Fellow and Professional Faculty Member at the University of California Berkeley where she focused her research on processes for design, innovation and systems thinking and developed the tools for experimental education. Matt is the Founding Director of the SEI's Emerging Technology Center and an Adjunct Associate Professor at CMU's Institute for Software Research. He leads the ETC in applying advanced technologies in artificial intelligence, computing and human-machine interaction to critical defense and national security missions. Before joining CMU, Dr. Gaston was the Director of Research for the Business Unit of General Dynamics C4 systems and spent nearly 10 years at the National Security Agency developing and applying mission-focused technologies. Frank serves as a Program Development Manager for the SEI's ETC. He's responsible for the business and portfolio management and customer relationship activities with intention to the DOD and IC. Prior to joining CMU, Mr. Redner retired from the U.S. Air Force serving as a Communications and Information Officer in a variety of assignments to include tours and Joint Staff, Air Staff, HCOM, Wing and Squadron levels. Now I'd like to turn it over to Dr. Rachel Dzombak. Rachel, good afternoon. All yours.

Rachel Dzombak: Thanks so much, Shane. And thank you, everyone, for being here today for this important conversation. We are so excited to talk with you about establishing the AI Engineering Discipline in the national initiative that we here at CMU are leading to build an ecosystem around this emergent field of AI engineering. Over the course of the next hour our hope is that you will learn a little bit about our perspective on the motivation for creating this field of AI engineering, to get too know a little bit about the current state and research directions in our three focal areas that help ground our conversations around AI engineering. And really, the focus of what we want to talk about with all of you is how you can get involved, how you can be a part of this national initiative and help grow it right alongside us. I'm so excited to be joined by my colleagues Frank and Matt. You'll hear from them shortly. And I know that they share our excitement in just starting this conversation. It's been a long time coming and we've been doing

a lot of work behind the scenes and we're excited to talk about it publicly here today with all of you. So why do we need this field of AI engineering? Well, right now, there's much attention from industry and academia focused on artificial intelligence. The development of AI technology and implementation of AI technology are often two different stories. Much of the current AI solutions that exist today are speedily deployed. They're tested in controlled environments with perfectly curated data sets. And they're difficult to replicate, verify and validate. What we want to do with this field is think about how do we progress from individual tools? How do we move past just capability developments to build and implement human-centered robust and secure and scalable systems, systems that really meet the outcomes that they intend to achieve? And so towards that end we want to start off with a question to all of you. We've been doing these conversations with stakeholders and asking them the same question. So to put our audience a little bit here on the spot, I would love just to ask you this question and Shane's going to put it in the chat for us of what is your level of understanding on how to leverage AI for real world outcomes. What do you think your level of understanding is on how to leverage AI for real world outcomes? Maybe you're at a 1, you're joining this because you have an interest in AI but you don't actually know what it is, you don't know how to distill AI from other technologies. Maybe you're at a 2, you know what AI is but you have no idea how it will be implemented. "Good luck with that," maybe you're thinking in your head. Maybe you're a 3, you have a lot of guesses on how to leverage AI but you're not actually sure. Maybe you're a 4, you have some direct experience. Or maybe you're a 5, you currently are working on this. You know exactly how to do this. This is your world and you have a lot of answers. So we're going to wait a couple of minutes or a couple of seconds, not minutes, don't worry, to see what answers roll in here. I'm excited to see that we have attendees from the United Kingdom, from Baltimore, from Nashville, across the U.S. That's fantastic. So we'll give it a couple of seconds just to see some answers roll in here. And I have to say, I've asked this to a lot of audiences and it's always great to see if people are willing to be honest. There's often a moment where people think, "Oh, I should put the answer that I think you want to hear." But it's always good to gauge really where people's understanding is. So we have some practicing data scientists designing and implementing some of the machine learning algorithms in the audience. So people at a 5. People at 2s. Spanning it. Just waiting here for a few more answers to roll in. 3.5. I'm glad to see no one yet is at a 1 which shows that we have a great audience for this conversation here today. But definitely, the results that are coming in are pretty typical. We see that a lot of people are at the level of about between 2 and 4, you know, most people falling into that category of 3, a couple of people saying, "You know, I'm working to implement AI systems right now. I have some tools and best practices behind me." But a lot of people are still at the stage of a 2, of, you know, I know I'm following the developments and emergence of AI technology and I'm thinking about the potential that it has to create impact. But I'm not exactly sure how to implement it or implementing it

seems really hard. And so that is why we're having this conversation today, too talk a little bit about that challenge and how do we build tools. And so towards that goal I wanted to just kick it over to Matt and ask a little bit about, you know, why is implementing AI harder than other traditional software systems? What makes it different? What's different about AI that leads us to ask this question of how to implement it for real world outcomes?

Matt Gaston: Well, thanks, Rachel, and thanks to everyone that's joining today. We're really excited to be doing this. I also want to say right upfront thanks to the SEI team behind the scenes that are making this all work for us today. So the question of why is AI different or why does AI introduce interesting challenges into the building of systems, in particular, production systems? I'd like to start, you know, we're from the Software Engineering Institute so we'll-- I'll start the answer by talking about software, traditional software systems. In those systems, you have teams of developers usually, maybe a distributed team of developers, writing down explicit instructions, explicit lines of code to build up the functionality. And that functionality can be very rich and complex, but it is built by these teams of developers over time. You can analyze that code that those developers create. You can find and debut and through that analysis understand how to fix that code or add new features to that code. And for the most part, I'll call traditional software systems determinist. Deterministic meaning I know the behavior of the system. I can provide unit tests and with certainty know that the flow of the program or the software system is working the way we want it to work. In AI, things are a little bit different. And I want to be careful here and not just say AI broadly, I want to be specific and say today in 2021, the most common way to get at an AI capability is with machine learning. Well, usually with deep learning or what people would call modern machine learning. And the way that works is there's no programmer writing down the instructions. There's usually a team of data engineers that are creating some data that's going to be used to train the system. So there are these machine learning algorithms that look at the data, that analyze the data themselves and produce some set of functionality. That functionality is based on math and complicated algorithms and the information that's in that data set. It's also really important to be careful that that data set is representative of the problem that you want and isn't riddled with any issues like bias or privacy concerns. And so the data itself actually drives the behavior of these AI capabilities that get produced via machine learning technologies. And once you have that technology, let's call it an inference engine or a model, that you can embed into another system, it's actually fairly opaque, right. It's a lot of numbers, weights, connections between different components of that neural network and it's very hard to analyze. And so unlike traditional software systems it's much more difficult to debug to understand where there might be failure modes or edge cases that you need to address. And then the last thing I think that's important about AI systems to mention right upfront is that they're not as predictable as a traditional

software system. I called traditional software systems deterministic. In most modern machine learning applications and systems, there is some amount of uncertainty, some amount of unpredictability inside of those. And so for the system developer, you have to think about how to use that uncertainty smartly throughout the entire system. And so that's why I think AI introduces both great new capabilities but also particular challenges when it comes down to engineering AI systems.

Rachel Dzombak: Thanks, Matt. And it's exactly that. Because of these challenges, because of the challenges that are unique to the field of artificial intelligence is why we are aiming to create this field of AI engineering, to start to think about finding those answers and sharing them so that we can move towards implementation of AI systems in a way that's rigorous, in a way that puts thought into all of those implementation steps and shares best practices with others. And so in just trying to start this field we've been engaging in a lot of stakeholder conversations and I was talking to someone about how they are implementing AI systems and kind of what are the practices behind their work that they're engaging in right now. And they said to me, "You know, Rachel, there's no book of spells, there's just magic." And I sat back and reflected on that or a second and then I thought, "What a wholly unsatisfying answer." Because first and foremost, AI is not magic, and when we talk about AI as magic, it creates a barrier from the diverse number of people and stakeholders that we need to engage in this conversation because of all of the reasons that Matt just laid out. Beyond that, it's not just that this is something that could be done by a select few. And we here at the SCI, we have a hypothesis that there are best practices, processes, tools, frameworks that can improve the deployment of AI systems over time and enable trust and confidence. We don't have what all of those best practices, processes are today, but our goal through this national initiative is to aggregate them, to put them into conversation and to share them across the community very much in a way that can enable what we want, which is this implementation towards mission outcomes and achieving of mission outcomes. So how we see the field of AI engineering today, our current definition is that AI engineering is a field of research and practice that integrates the principles of software engineering, computer science, systems, human-centric design too create and implement AI systems in accordance with human needs for mission outcomes. And I say that's our current definition because we view that this field will evolve as AI technology is evolving and we want to be very mindful of how we keep learning and evolving our notions of this field alongside the technology itself. And so towards that end, you know, when we're trying to define this field we're trying-- we're using the problem context of defense and national security to guide our definition. We're doing that intentionally. People have asked us, "Are you defining the field? Are you defining AI engineering for defense and national security?" And we really view it as the former. We're trying to put thought leadership into what is this field, what are those practices, tools and

frameworks but using this problem context. And Matt, I'd love for you to add a little bit about why national security. Why does it-- What is relevant about the national security problem space for AI engineering that makes it rich for the definition of this field?

Matt Gaston: Of course. Well, first, the Software Engineering Institute is a defense, the defense software research lab. And so we have a natural focus on the special needs of the defense and national security community. But there are some interesting fairly unique challenges that I think exist in the defense and national security space when it comes to the adoption and deployment of AI technologies. One simple one is that the concentration of talent compared to say a big tech company, and you can pick your favorite one in your mind, is just not the same in the government. There are plenty of really, really-- really, really smart, really, really ingenious, really, really innovative engineers and scientists in DOD. But again, the concentration of talent that understands AI and how AI will work and affect mission is not as concentrated. So that's a big one. Work force we would call that. Second, there is a variety of different challenges, different modalities, mission sets that are really spread across a very broad spectrum where if you look at a big tech company, a big internet company, they're pretty focused on a single thing, right? Recommending movies or solving search queries or serving advertisements that are relevant to a particular domain. Where in the defense and national security space, the problems vary across a much bigger spectrum. There's also issues with data in the defense and national security space. You don't always have all of the data you might need to train a system. It might not be readily available. You might not even be able to create it. And so a lot of ideas around how to build these systems where you're data impoverished or you don't necessarily have a way to get to scalable data sets. And there's many other challenges. There's the mission criticality or life criticality of applications. There's the acquisition process itself and how that process can handle understanding requirements for AI capabilities and also all the way through testing, deployment and even modernization. And so there's lots of challenges. Of course there's much to be learned from what industry has done over the last decade or two in using AI technologies. So we don't think it's a completely unique problem set for defense and national security, but we do think there are some special issues that have to be addressed.

Rachel Dzombak: Absolutely. And certainly where we started this conversation about current AI developments being in controlled settings when difficult to replicate and national security, that's the opposite where we're dealing with really complex problem spaces, siloed data as Matt mentioned, all these challenges that need to be overcome. And that's where the engineering part comes in of how do we start to define these best practices around these challenges and our thought is if we can solve that in the national security space, there's lots of lessons to be learned for other industries that are facing analogous challenges. So a national initiative pushes

us towards this goal. In partnership with all of you we hope to create and expand the AI engineering body of knowledge, tools and practices. We hope to engage in workforce development, in digital transformation initiatives to put these practices into organizations and to help deploy AI in a rigorous way, disciplined way. We're really launching this national initiative to first and foremost formalize the field of AI engineering to just be that group that is helping to convene perspectives, practices, tools across a broader community, but also to establish an ecosystem. Because we know that these conversations are happening, we've been having them in pockets with our partners and we want to centralize those conversations. We want to be able to share lessons learned across different industries within the defense and national security space to make sure that we are moving forward, moving towards this goal because it's a really hard problem. And so that is what this national initiative, what we're hoping to do is to have a committed partnership towards evolving that ecosystem and really to develop this pipeline of promising AI engineering solutions and tools that can be used and adopted by a variety of organizations. And one of our big motivations has come from the national security stakeholders that we work with every single day. And those motivations were recently I thought very well-articulated through the National Security Commission on Artificial Intelligence Report that they released on March 1st. One of the quotes from the report that I loved was, is up here on the screen, and it's that "This new era of competition promises to change the world we live in and how we live within it. We can either shape the change to come or be swept along by it." And certainly I think with this field of AI engineering, if we can define practices that help with the support deployment, we can be proactive in thinking about what this technology can do, how it can help us achieve our mission goals. Versus being in a state where we're reactive to it. "Oh, we know we need this. How can we integrate it really quickly?" AI engineering is what's going to support us to be in that proactive stance and help shape the change to come. And so towards that end, Frank, I'd love to turn it over to you and just ask a little bit about your motivations. What did you hear from our stakeholders that motivated this national initiative and our involvement in it?

Frank Redner: Well, first off, good afternoon, Rachel, Matt and all. I'm so glad to be here with you all. And Rachel, thank you for the question. I think there are many differing motivations out there for this initiative. First, there is a recognition that there is so much we don't know about the dynamic and adaptive nature of AI systems which Matt spoke to earlier. So much we don't know meaning so much we have to learn. There, too, is a realization that there is lots of research and development going on in many areas from developing effective methods for human-machine collaboration, understanding and addressing the ethical, legal and societal issues, ensuring the safety and security of AI systems, to actually measuring and evaluating those systems through standards and benchmarks. That's just to name a few of the R&D activities ongoing and this,

again, was pointed out earlier. No one, no one is aggregating and synthesizing those lessons learned for the broader community. And then finally, there is a sense that we're embarking on a much needed journey. There is excitement and genuine interest from individuals and organizations in uniting around a movement that really seeks to maintain our competitive advantage in AI and get us to the justified confidence that's prescribed in the final report from the National Security Commission on AI. So Rachel, those are just a couple of the motivating factors that I've heard.

Rachel Dzombak: Thanks, Frank. And Matt, I'd love to turn it over to you if you have anything to add there.

Matt Gaston: Yeah, so I've seen this quote and heard people use this quote very recently and it comes largely from industry, that 85 percent of machine learning projects in companies fail. They never make it to a mission or a business capability. They never make it into production. Eighty-five percent is a staggering number. And if we add the additional challenges that the defense and national security community might bring to these problems, there's a huge concern. There's a huge concern about how to take advantage of these amazing capabilities of modern machine learning and other AI technologies. And so if 85 percent of industry projects are failing, we want to understand why. We want to understand what has been tried and what doesn't work. Why these things fail, why they stop working over a period of time, which is a very common problem. You might train a particular machine learning model on a data set and it may work in production for a week or a month but then it stops working. How do we understand these problems and how do we put the engineering processes in place to detect these problems, mitigate them and maybe evolve our systems over time? And so I think our big motivation, and I want to say right now I think we'll probably talk about this more, we don't think we can do this alone. This is a movement of practitioners, of people sharing war stories, best practices, capabilities, tools. But we want to create a body of knowledge, organize a body of knowledge, a living and breathing body of knowledge that helps people push that number from 85 to 75 to 65 and maybe even better than that at some point. And so I think that's the real motivation is how do we have more of the investments into building of these technologies be successful investments, whether they're in industry or in the defense sector.

Rachel Dzombak: Thanks, Matt. And towards that end, I just want to build on both that you know, we don't think we can do it alone nor do we think that we are the only ones working to do this. Our hope is that our community at large, all of you who are kind of joining us today, can help to surface where else is this being done, who else is convening and articulating these best practices. Because we have a community over competition mindset, wanting to be in a place

that we can-- this problem is too big to be solved by any one discipline, any one group of individuals. And I think the more we can pull these together, we will be, towards Matt's goal there of reducing those numbers of failures, we'll be more successful in our attempts but also be producing a better system with it that achieves this notion of assured AI or systems that we can feel confident in that are resilient over time. Encourage you to put your questions in the chat. We're going to get to those just here in a few minutes. So if you have questions about what we've talked about so far, feel free to throw those in and we'll get to those in just a couple of minutes here. And so I just wanted to be transparent about what our goals are. And honestly, it's a really hard thing to say how are we going to put a stake in the ground around the discipline. And so to bring our internal team together, here are some of the things that we're working to do. Here's how we see our goals in bringing this national initiative forward. We view it as a responsibility to help to define the discipline, to articulate those processes and iterate them so that we can frame and solve AI engineering problems. We want, it's not enough just to be able to solve those problems, we also have to be able to communicate what we did and how we did it so that we can bring others along with us, recognizing we don't want this to be a skill set held by only a select few, that in order for AI to reach its potential it needs to be broadly shared and broadly practiced. We want to cultivate the workforce and that connects to our first goal there of if we can articulate these practices, then we can share them with others. We know that the workforce needs to be continuing to evolve with the technology itself. And we have a large outreach mission hoping that we can share these tools and practices, help adopt them and it's not that we're just going to be speaking to data scientists and machine learning researchers, there is a large number of stakeholders that are needed to learn these practices so that we can ask the right questions, acquire the right technology, bring data sets together in a unique way. And so part of our goal is to help articulate who is involved, who should be part of that workforce. And then how do we get them to where they need to be. We want to help steer the ship, to identify and direct energy of other researchers towards critical AI engineering growth areas. And that is not because we think we have a unique eye on special questions, but it's really because of what Frank described as someone being this aggregator. We want to be able to be seen across the field of what people are working on as a means of driving work forward in areas that maybe aren't as fundable at the current moment or aren't getting as much attention, so that we're moving these different aspects together in parallel. At the same time, we at the SEI, you know, we also are a research organization and so we are going to be practicing what we preach. You know, we have current work going on where we are doing reflections to surface okay, what were those practices that we did? In our own research efforts, are we doing good AI engineering? What does good even mean? What does that look like? How do we capture those lessons learned over time? And starting to put that structure in place to surface from our own work what AI engineering would look like-- looks like. And the last goal which is the most

important is to create ecosystem opportunities for researchers and stakeholders to share, learn and evolve the field. That hopefully will look like publication venues. It will look like a conference where we can bring our community together once we're allowed to be back in person together. Or doing it online like we are today, but making sure that we are connecting the groups that we have access to. And the SEI is really uniquely positioned to be a convener in this way because of the ecosystem we sit within and the partners that we have as well as our connections to industry and academia. And so before we move forward, Frank, I want you to go a little bit deeper into why it makes sense to pursue these goals here at the SEI. I just articulated a few reasons but would love for you to go a little bit deeper into that.

Frank Redner: Yeah. So, Rachel, I'm glad you asked that question why SEI. Because as I've heard that a few times while conducting our one on one practitioner interviews and even sitting down one on one with senior leaders. "Why the SEI?" So personally, I feel we're uniquely positioned to really unite a community around AI engineering for a number of reasons. One is our mission as it's been for nearly 40 years now is to support the nation's defense by advancing the science technologies and practices needed to acquire or develop, operate and sustain software systems that are innovative, affordable, trustworthy and enduring. I'm glad I got that out. I read it word for word. But that's in our sponsoring agreement. And what we view AI as being is a special type of software. So that's reason number one. I think reason number two is we have great connections with academia to include Carnegie Mellon University, which provides us access to world class faculty and staff. So for example, Carnegie Mellon University School of Computer Science is ranked number one in the nation, but to be fair, they are tied with Stanford, MIT and University of Cal Berkeley. But if you rank them in alphabetical order, CMU comes out on top. CMU also has the number one AI program in the country and we have regular interactions with the Human-Computer Interaction Institute, the Institute for Software Research, the Robotics Institute and the National Robotics Engineering Center, so we've got a lot of world class talent that we can tap into. Again, not only here at CMU, but across the country. And I think third, we're an independent government research lab serving in the role of a trusted third party advisor expected to provide unbiased input. And even though we have these three reasons and I think we're uniquely positioned as we are, as we've heard all along, this doesn't guarantee success. An endeavor such as this is going to take a community of dedicated partners to realize that evolution of the discipline focused on delivering systems that are scalable, robust and secure, human-centered, providing for expected mission and business outcomes. So I hope I gave you enough, but that's the way I see our position and that we're uniquely positioned to carry out this initiative.

Rachel Dzombak: Thanks so much, Frank. I particularly liked the emphasis on where we come in the alphabet. Definitely part of our strategy here overall in rolling out this discipline. Before we move forward we have a couple of questions from the audience that Matt, I'd love to turn over to you. And so, one of the ones that just came in is, "Why do you think the focus across industry is on custom written AI rather than commercial off the shelf modules? Or embedded integrated AI available in commercial off the shelf products?"

Matt Gaston: Yeah. I think I can give a fairly simple and short answer to this question. There's definitely a place for COTS models or modules that have been built and are available for purchase or use from commercial or open source. But if you look at any-- at many machine learning applications today, someone, the developer is maybe grabbing a model that has already been trained for some other problem or some related problem but there's always a need to customize. And in some cases, you have to customize it with a whole new data set. And so COTS solutions in this space, again, may be practical when there's a really good fit between the business or mission problem and the solution that that COTS product is providing, but almost always in trying to deploy these technologies they've got to be customized to the particular use case and the particular operational context that the systems will be deployed in. And so there's a role for COTS but also there's a role for customization. They're going to have to both happen.

Rachel Dzombak: Absolutely. We're definitely going to need a portfolio of solutions towards building these systems. But another question just came in of "Why do you think that the--" Oh, sorry. I was looking at the wrong one. <laughs> "There's been a lot of lessons learned with CMMI high maturity experiences in regard to data modeling and statistical analysis. In what ways are these experiences being leveraged by the AI community?

Matt Gaston: Well, certainly, you know, we're from the Software Engineering Institute, which was the originator of CMMI and the legacy of Hua Tung tree [ph?] and all of the great people that worked on that. And there's much to learn, right. We want to build on that legacy and the knowledge that was created through that process. But we also need to look at best practices, tools, methodologies that are adaptive with the pace of technology innovation in this AI space. Not to say that maturity models aren't, but again, you know, the highest level of maturity model is usually adaptive and tailored to a specific application or a specific domain. And so much to be learned from those things. There was a question earlier from Susan Brown about data maturity models. Not endorsing any particular maturity model, but certainly data readiness, an organization's data readiness is imperative to being successful in adopting AI technology.

Rachel Dzombak: One more that just came into the chat which is more of a comment but I'll flip it into a question. Martin said, "From my experience, AI engineering and object engineering are not that much different. For both you need examples and lots of them." And Matt, you started talking about this a little bit earlier on the need for AI engineering. Could you talk about kind of what's so hard about gathering those lots of examples in that way?

Matt Gaston: Well, specifically, labeling data can be very expensive. Right. Just getting a pile of data is maybe not all that difficult but getting data into a format where the features in the data sets are usable, that you've analyzed the data set to make sure it's not riddled with problems like bias or other things. And then labeling. Labeling is actually, I think, one of the biggest challenges with data sets. A lot of data, certainly in the defense and national security space, might not come labeled. It might not come with the sort of answers attached to it. And in that case you have to get creative. You maybe have to simulate things or other approaches to creating a data set that can be used to train a modern machine learning model. There's also lots of work sort of in the academic space on what to do when you don't have labels. How do you do unsupervised learning? How do you organize your data set and train a model that can find the differences and start to understand the different semantic relationships that might be in a data set. So there's much to do. Data readiness is again I think a huge part of what we're after and we need to collect all of the lessons learned. There's some great work at Stanford on a project called Snorkel which is about using traditional software, little rule-based programs to automatically label a data set. But you use lots of different rule-based programs to do that and then you train your model on the aggregate data that comes out of those smaller more rule-based, more predictable codes.

Rachel Dzombak: Thanks, Matt. I'm going to keep us moving forward a little bit. A couple more questions rolling in and we'll make more time for those as we go forward. And so, you know, for us I think one of the central challenges was where does one start in terms of defining a discipline and how do you even start to walk into that space. And we started by doing what we're doing right now but a little bit more in proximity than we are currently of talking to our stakeholders. We typically don't like to be separated by a Zoom room and this webinar format. But we started having a lot of conversations with our stakeholders to try to understand exactly what we've bene mentioning of what makes implementation of AI systems so challenging. And I think Shane's going to put into the chat here a report that came out of one of those workshops we had with our stakeholders and summarize those. And but from those conversations led to what we call our AI Engineering Pillars, which are really grounding focal areas to help us move forward in this large ambiguous space. And so for us we have three different pillars that are guiding our work: scalable AI, robust and secure AI and human-centered. Scalable is thinking

about the ability of data models' algorithms and computing infrastructure to accommodate the speed, size and complexity and complex needs of training and inference tasks dictated by the mission. For human-centered AI, it's AI designed to work with and for people, encompassing lots of topics of ethics and bias and how do you actually build trust between human and machine teams. On robust and secure, that's where we're focused on how do we get systems to reliably operate at expected levels of performance even in the face of uncertainty or in the presence of danger or threat. And across those areas we're starting to articulate what are those open questions, what is the state of the art, who's working in those spaces, and pulling those lessons together. And Matt, I'd love to turn it over to you to talk a little bit about what do you see as some of the challenges within those pillars and across them right now?

Matt Gaston: So, Rachel, you know that's a massive question and we could probably do, you know, an hour or two on each of these pillars and all of the challenges and opportunities as well as, you know, best current practices that align with these different pillars. So I'll just say a little bit about each and maybe I'll go in reverse order. On the outset of this initiative when we got started, I will actually admit that I was surprised that human-centered AI was one of the three pillars. Of course that's very important, but it became through the conversations with our stakeholders and the folks we were talking with, it became central and I think that's actually a great thing. It's the right thing. We've got to figure out ways to design these systems with their human counterparts, with their users, with their partners, human partners in mind. A key scientific challenge for an AI system is what's called value alignment. And what this is is how do you align the objectives that are given to the AI system, maybe the optimization function that's given to a machine learning system. How do you align that with the objectives of the human counterparts? And this is a hard scientific question. It's an active area of research. There's amazing work happening in this area. But even more important than maybe getting alignment right at the creation of these systems is maintaining alignment over time with these systems. Because of course, the needs and objectives of humans or teams of people can change over time. And if those are changing, how can the AI system adapt and evolve with it? Of course, also aligned in the human-centered pillar is the AI components of these systems need to provide sort of information or confidence or even trust. Again, a very active area of scientific research, what does trust mean in these systems, how do you signal trust, how do you produce evidence that will lead to trust between both the system and the humans. The second pillar, robust and secure, again going in reverse order from the slide, this is a huge one. It's really test evaluation verification and validation. How do we know that these systems work? How do we know that they-- why they're going to fail? When they're going to fail? And what to do about it? Really important in this is that again, you need to monitor these systems over their life cycle. You can't just test once, certify them for operations and then let them go off and operate

because there is some unpredictability, there is some uncertainty. There can even be drift in the environment that they're operating in or changes in context. A second big thing here is that these systems in part because we don't fully understand exactly how they're working or they are massive correlation engines, they do introduce novel vulnerability or attack surfaces. They can be manipulated in intentional and unintentional ways. And so building up some science around that that leads to engineering best practices and tools will be critically important over time, right, especially as new algorithms, new models are being generated every couple of weeks as it turns out. So robust and secure is really about knowing that these systems work and that they continue to work over time. Another really important piece of robust and secure is that most of the time when a machine learning model is built, it's evaluated at model creation time and it's typically evaluated for accuracy, how well does it perform on the particular task. Our thoughts on AI engineering today are that we've got to get beyond the accuracy. We can't limit ourselves to just focusing on model evaluation, model accuracy. We need to create metrics, business metric, business use case metrics or operational use case metrics or mission use case metrics that the models should be evaluated on right at their creation time, not just accuracy. And then the last area or the first area, scalable AI. Rachel said it. In defense and national security, but also in many applications across a whole variety of domains, these technologies still need to scale. They still need to scale to the size, speed and complexity of the problems that are out there in the world. Another great quote that I heard is that some of the, you know, most recent models that have been published by places like Open AI have cost $4 or $5 million dollars to train them. If you imagine trying to run a business where you're having to train models all the time but each one of them costs $4 or $5 million dollars, that could get quite expensive. And so how do we scale the training? How do we push these types of technologies to edge computing, be it IOT or other devices? There's certainly applications in the space domain that are quite interesting where you might be sort of disconnected from those models. So there's lots of challenges in scalable AI alongside speed and complexity. The other part of scalable AI is really enterprise scalability. How do we make these capabilities available in a responsible way to a broad community of stakeholders in an enterprise, be it a military service or a business that wants to apply these types of technologies, that wants to take advantage of these types of technologies? How do you do enterprise scalability in a responsible, repeatable and manageable way? And so those are the three pillars. There's of course common threads across all of them. One of them is this idea of testing across the life cycle of an AI system. And notice that the life cycle necessarily is a spiral or a helix or whatever you want to describe it as, as the system's going to evolve over time. So take the current DevOps process that many people are using and apply sort of monitoring and testing across that when it comes to AI.

Rachel Dzombak: Thanks, Matt. I know that that was an impossibly hard question but you did a great job diving into it. And certainly that's not all we have to say on those topic areas. And hopefully there's a comment in the chat there with a link to sign up for our newsletter. We've been working on kind of three papers that have been vetted by our government partners as well as colleagues and excited to drop those very soon. And so if you're signed up for that newsletter, you'll be the first to gain access to those white papers on each area as well as the summary across all three once we have those ready to go here in the next few weeks. I want to touch on a couple of the other comments that were in the chat. So Duane asked, "Will SEI's effort include code repositories, compute resources, other resources?" Certainly the newsletter is going to be a great place to find access to those over time. We're exploring lots of different mechanisms of community building right now, thinking about how we can share resources across this ecosystem that we are aiming to create. And our hope is that over time we'll be able to share more of those across our community. And certainly our access to our website will be a great place to find those. Another question, there was a question about, "How far is human-centered AI amongst neural diverse people as well as kind of what are your thoughts on using AI to recognize disinformation in social media?" And I wanted to say on both of those, hugely important questions, and that's exactly what we're trying to do with these pillars is to spur work in really interesting questions. We have just been having a conversation about a potential project on how does data labeling look different when done by neurotypical versus neurodiverse populations and how does that lead to a different set of outcomes? And so exactly these questions that are surfacing in the chat, that's the type of work we want to seed through this national initiative to help address these hard problems. And so, what I want to do next is just move us to our community building which is, I think is hopefully why everyone's here today to think about how we can get involved. And so how do we envision the community getting involved? We have a couple of different mechanisms here that we would love to see participation in. The first is being an advocate for AI engineering. We need partners. We need partners, we need that community building behind us right now. And we can't, you know, as much as I would love to stand in a Zoom room and talk every single day to gather support, we have a lot of work to do and so we can't do that part alone. We need advocates across industry, across government to be starting to ask how are we thinking about establishing best practices around AI deployment. How are we thinking about AI engineering? And pushing the conversation to be less about what's shiny and new and emergent to how are we doing this in a rigorous way that instills trust and confidence in the users and achieves that mission outcome. We want to do collaboration to build the discipline. If you have a research project you're working on that you think is relevant, we'd love to hear from you. We'd love to do an active partnership with you in some way. And that will look like having a conversation with us to think about where and how we can connect. Over time our hope is to open up funding opportunities to our

community to spur joint research, to spur independent research on these critical topic areas. And so towards that end, we'd love also to think about for funding agencies to help support our research agenda, to help put funding dollars, support, use cases behind these research initiatives and to help to drive them forward, help to make sure that we're able to do the research that is so critical to be done in this field, both in terms of surfacing what those best practices are as well as making sure that we are moving forward and making steps with the systems that we are trying to implement today to ensure that they are engineered in a way that meets these goals of being human-centered, robust and secure and scalable. And so we're going to get to questions and answers. This is my last slide. It's just ending on these really concrete opportunities to engage. If you want to meet to talk about partnership opportunities, we have set up a weekly office hour and you can find it at the link that is up there. The slides will also be sent out afterwards so that you'll have it. So you can meet with us to think about who's the right person to connect you with. What could a research partnership look like? How can we amplify work that you're doing in your own organizations? We're hoping that through those conversations we'll be able to get to know stakeholders who are in our community a little bit better and think about what does a concrete partnership look like. Point us to your research and work. If you are sitting there listening and thinking, "Oh, they really should be looking at what Organization X is doing," please tell us. There's an email address there at the bottom of the slide. If you reach out, we will be able to definitely-- We would love to take a look at what you know about towards this initiative that we should be paying attention to. I already mentioned signing up for the newsletter. That will be the best place to access information about new events, opportunities, ways to connect and research that is going on in the space of AI engineering. And the first thing that we're aiming to release is our white papers on these different pillar areas making sure that we're just putting out what is the state of the art in those different pillar areas as well as what are open questions that we see that we would love to get all of you working on because we certainly can't solve all of those ourselves. Frank, anything you want to add on opportunities to engage? Anything I missed?

Frank Redner: No. You know, we've done this before. We've worked with small business, bought [ph?] small businesses, did partners, academia, state and local government, federal government and of course DOD and the IC community. So we've done this before, not on this scale. So this is doable. This is something we've not done of this magnitude, but we are looking forward to sitting down with each and every one of you and finding how you can participate in this initiative.

Rachel Dzombak: Matt, anything you would like to add on that?

Matt Gaston: No. Just to reemphasize that we think of this as a movement and we're looking to convene the community to bring people together to find great ideas wherever they are. So we look forward to working with everyone out there.

Rachel Dzombak: Great. So there's another question in the chat of, "Will you offer education opportunities for the DOD and other U.S. government employees?" I would love to say definitely it's something we are thinking about what's the best mechanism to do it. So yes, we are thinking about education opportunities. We're currently trying to unpack what that looks like. Could it be a specialized, you know, one year program by which we educate people? Is it a training opportunity? We're doing some work right now in the intelligence community helping to do workforce development initiatives. But if there's something that you have in mind for your organization or you could really benefit from training in XYZ, please do reach out to us. We'd love to talk about those specific opportunities and how do we help you to achieve the goals you have, whether it's us delivering that education or connecting you with one of our ecosystem partners who we also are working with to think about what does education in the space of AI engineering really look like there. Another audience question that I would love to throw over to Matt is, and trust me, it's not as hard as my question on what can you-- outline the initial challenges across all of our pillar areas in two minutes or less. But, "Could you compare your AI engineering campaign to the MLOps campaign of other areas?"

Matt Gaston: Sure. I think MLOps is a piece of what we're thinking when it comes to AI engineering. I'm reminded of a course at Stanford right now, I think it was first offered this spring, on productionizing ML models. Chip Huyen I think is the person's name that's teaching the course. I've seen a little bit of it. It's really great stuff. I'm glad that we're starting to see sort of practical engineering education happening at places like Stanford and other great universities. So MLOps is really a piece of the puzzle, but maybe doesn't go far enough. It doesn't encompass enough of what is needed for all of AI engineering. MLOps is really about creating models and getting models into production. There's a lot more that goes on. There's new techniques for testing and evaluation that I've already mentioned. MLOps doesn't necessarily capture one area that I failed to mention for Rachel's hard question, which is in human-centered AI, one of the big focus areas for us is on ethics. Not just writing down principles of AI ethics for a particular problem domain, but really, how do you implement ethical mechanisms in these systems. And again, MLOps is a piece of the story here but how do we build these overall systems and take into consideration things like ethics and what to do about them. In fact, we do have an AI ethics checklist, design checklist, that's out there on the SEI website. Shane might have it handy to drop it into the chat, I'm not sure. But we're starting to create and accumulate artifacts like that that can help teams build these systems and take into

account ideas like ethics. So to answer the question succinctly, MLOps is a piece of the answer and we're excited about what's going on in that space.

Rachel Dzombak: Thanks, Matt. There was-- Or Frank, did you want to hop in? You looked like you had a--

Frank Redner: Yeah. There was a question in the chat that, "The OND has recently chartered an AI Taskforce and who would be a good point of contact for that group to contact to begin any collaboration?" And that would be either me or yourself, Rachel. And that's why we set up the office hour. So just reach out to us and we'd be glad to sit down and have a chat and talk further about how you can participate.

Rachel Dzombak: Thanks, Frank. There was a question a little bit earlier or two that we had skipped over that I'd love to go back to. Matt, the first question was for you of, "Could you explain data readiness in a little bit more detail and what that looks like?"

Matt Gaston: Yeah, I actually just had this conversation recently with a friend at an accounting firm, an accounting and auditing firm. And so what data readiness means is really what's your organizational approach to data? The firm that I talked to hadn't really thought a whole lot about what their strategy was with data. They had a lot of data that they could be using, that they could be accumulating, that they could be starting to, you know, build small AI capabilities with. But you've got to get organized and there's different levels of getting organized, right, from how you're storing and structuring that data, how you're sharing it, how you're protecting it. So data readiness is how good are you and how good are the practices and processes are for dealing with data, accumulating data, managing, sharing and maintaining data sets. Because after all, at least in modern machine learning, data is really the driving force behind these technologies.

Rachel Dzombak: Thanks, Matt. There was another question earlier that I'll take and I view it as a hard question, so just to be fair to Matt. Where there was a question of, "How can academia help to ensure-- Or how can academia ensure the safeguarding of national security information?" And I don't think we can guarantee that academia can ensure that. And certainly, we wouldn't make the claim that we know the answer to that question. But our hope is that by doing, engaging in this national initiative, by bringing people together, we can make strides towards it. I think that we have a competitive advantage when we can surface the perspectives of the many and start to engage in those debates, discussions to move this conversation forward. And certainly I think our pillar areas are means of moving forward work and conversation and debate, all of those things, in these areas of what does it actually mean to be

human-centered? What does success look like? How do you measure towards that? There are a lot of open questions right now that are just sitting there. They're sitting there and we want to spur work in those areas, same with in scalable; same with in robust and secure. And so our hope is that if we can start to answer some of those questions, even in small ways, we'll be moving towards that notion of assurance in AI systems and helping to be more confident as we deploy these systems. And so with that, I want to just give us a second to close up here recognizing we're almost at time. And so I'll just give us a minute to do some wrap up comments and certainly as Frank mentioned, please do take advantage of these opportunities to engage, we'd love to hear from all of you, to think about how we can-- the role you can play in helping to move this discipline forward. And from my perspective, that means hearing from a broad number of stakeholders. I hope if you take one thing away from this it's that the moving AI engineering forward, it's not going to be done by just a couple of people. That we need all sorts of stakeholders, all sorts of roles: leader, ML researchers, data scientists, social scientists, anthropologists, clinical psychologists, all of these people who can join together to think about what these systems look like and where and how we implement them. And I hope that this field of AI engineering provides a space to have those conversations, those messy, hard conversations that are often hard to have in an organization, as a means of surfacing those best practices, figuring out ways to step into that ambiguity and to move us forward. And so, Frank, over to you. I would love to hear any final thoughts that you have.

Frank Redner: Yeah. I don't think I'm going to say anything new. I do believe that we're at an inflection point. Our competitive advantage is being challenged in many technology areas to include microelectronics and AI. And if we're to win the AI era, I think now is the time that we need to coalesce around this AI evolving and AI engineering discipline to enable justified confidence in AI for our national security and for high impact problems contexts. Now this will take a whole government, maybe a whole nation approach, and it will really take an innovation mindset. We have so much to learn about these systems, but we need input of many to make this progress. So please join the movement. I was at an NDI AI event earlier this month and the Chief Architect of the Air Force said in not so many words, we can either be disrupters or we can be disrupted. And I'll end on that note.

Rachel Dzombak: Thanks, Frank. Matt, over to you.

Matt Gaston: Well, thanks, Rachel and Frank. I can say this, you know, this hour flew by. It's really exciting. The engagement from the audience has been great. I'm so happy that there are people out there that want to be involved in this idea of AI engineering and what to do about it and help everyone, really, figure out how to build these systems in a responsible way to get to

the point where Frank keeps emphasizing, justified confidence in its capabilities. So I'm really, really excited about that and I'm really excited to see where things will go. I did want to take just a minute to say thank you to our sponsors. We have sponsorship from the Office of the Director of National Intelligence. We also have research partnerships in AI engineering with NGA, the National Geospatial-Intelligence Agency and Northrop Grumman. And we have an active collaboration partnership with the University of Maryland Applied Research Laboratory for Intelligence and Security, so-called ARLIS. And so we're really grateful for all of those partnerships and sponsorships. And of course, if anyone else is interested in getting involved, we'd love to hear from you. So thank you.

Rachel Dzombak: Thanks, Matt. Thank you, everyone, for joining today. We look forward to hearing from you. And with that, Shane, I'll turn it back over to you.

Shane McGraw: Rachel, Matt, Frank, great discussion today. Thank you very much for sharing your expertise. And I'd like to thank you all for attending today. Reminder, upon exiting, please hit the light button below the video window and share the archive if you found value today. Also you can subscribe to the SEI YouTube channel by clicking the SEI seal in the lower right corner of your video window. Lastly, join us for our next livestream which will be on April 28th and our topic will be announcing IEEE 2675 DevOps Standards to Build Reliable and Secure Systems. Registration information will be on our website and emailed out as well. Also we'll send out follow-up email today with a lot of the-- the location of today's slides and some of the next steps you can take from today's webcast that will go out shortly to everybody registered for today's event. Any questions, you can also always send anything to info@sei.cmu.edu. Thanks, everyone, and have a great day.