Carnegie Mellon University
Software Engineering Institute

SEI Podcasts

Conversations in Artificial Intelligence,
Cybersecurity, and Software Engineering

Best Practices and Lessons Learned in Standing Up an AISIRT

Featuring Lauren McIlvenny as Interviewed by Suzanne Miller

Welcome to the SEI Podcast Series, a production of the Carnegie Mellon University Software Engineering Institute. The SEI is a federally funded research and development center sponsored by the U.S. Department of Defense. A transcript of today's podcast is posted on the SEI website at sei.cmu.edu/podcasts.

Suzanne Miller: Hello and welcome to the SEI Podcast Series. My name is Suzanne Miller, and I am a principal researcher in the SEI Software Solutions Division. Today, I am excited to welcome Lauren McIlvenny, technical director of threat analysis in the SEI CERT Division. Lauren is joining me today to discuss the current AI threat landscape and how the SEI created an AI Security Incident Response Team [AISIRT]. Welcome, Lauren.

Lauren McIlvenny: Thanks, Suze. Glad to be here.

Suzanne: You are new to our podcast series, Lauren. Let's start off by having you tell us a little bit about yourself, what brought you to the SEI, and what is the coolest thing about the work that you do here.

Lauren: Those are all good questions. Like we said, I am Lauren McIlvenny. I am the technical director for threat analysis, so overseeing work from vulnerability research to malware discovery, and to this newly formed AI Security Incident Response Team. I have been at the SEI for about a year and a half. Prior to coming here, I was at John Hopkins University Applied Physics Lab, where I really grew up in their AI branch. It was super exciting to come here, like the birthplace of cybersecurity, and to bring that AI expertise, and then use that to really start forming this AI Security Incident Response Team. I mean, what better place to be than Carnegie Mellon [University] if you want to be in both cybersecurity and AI?

Suzanne: That is true. Let's start with the attack surface on this. I've learned all these new terms in talking to all of you folks in CERT: *threat analysis* and *attack surfaces* and things like that. But how do attacks on AI and the vulnerabilities that lead to them differ from those that we are used to seeing in non-AI systems, procedurally based software systems? And how are they the same? How are they different and how are they the same? I think that will help our audience understand why we need to actually have a special incident response team for AI.

Lauren: Yes. I think you are going to hear me say a lot more the same than different. Fundamentally, we here in CERT will tell you that *AI is software*. It is code that is running on some type of device, albeit usually a fancy one that is often connected to a network. A lot of what you are going to see, or that we have been seeing, are traditional cybersecurity vulnerabilities that are just going after this as software. But there are some differences, right? You will see different things like AI also has data in it. They are going to go after the data that was used to make that AI to train it. Models are executables, but they have things like weights. You can kind of think about them as configuration files in traditional software, but they are a place of attack. Really, I would say a lot more of the same than different but acknowledging that there are some differences that we are working to address. If you are going to say, *All right, Lauren, you just told me mostly the same then why form this, an AISIRT?* A lot of it is for awareness. We hear a lot, at least in industry, *AI is special*, and it does do a lot of really cool things. I am not going to argue that. But they still do need to get integrated into the cybersecurity community and think about things like vulnerability and secure by design. A lot of those principles we adhere to on the cybersecurity side for software, AI should also be adhering to, so trying to increase that level of awareness.

Suzanne: The way I am looking at it is, in the '80s we were becoming aware of software as a threat landscape. We had a lot of practices and a lot of software that was vulnerable because we weren't paying attention to that. What I am hearing you say is that we are kind of in that nascent stage with AI where everybody is building AI as fast as they can and doing all the cool things, and *Look what it can do here* and *Look what it can do there*, but not necessarily bringing with them all of the cybersecurity awareness and secure coding practices and things like that from procedural kinds of systems that also apply to AI. Is that a correct way of looking at it?

Lauren: Exactly. You just hit the nail on the head.

Suzanne: Even though it is a fundamental shift in technology, then we are really just dealing with… It is a little bit of *Groundhog Day*. Same stuff, different day, but the stuff, it may be a little sexier. It may have different parameters, like more data centricity, as you mentioned. But we need data security. We need model security. We need algorithm security, just like we needed library security and coding practices and no buffer overflows and all the things that we have come to learn and take for granted in procedural systems. We just don't have that list. We have a general list, but we don't have necessarily a specific list of what are the vulnerabilities that we need to pay attention to in AI. Is that fair to say?

Lauren: It is absolutely fair to say. When we talk about vulnerabilities in AI, even some of the things you mentioned, like libraries, are still here.

Suzanne: Sure. We just have a different character.

Lauren: Yes, absolutely. I will say that was one of the biggest things that I am finding doing this is, the vernacular isn't the same through the teams. So you will talk about a software stack. That is something completely different than what they mean on an AI stack often. It is a lot of translating between the two communities to get things kick-started.

Suzanne: Okay. That alone sort of offers the opportunity for an incident response team that helps to make those translations and that helps to make the AI folks aware of the vulnerabilities world. Then I am also assuming it also makes the cybersecurity folks aware of this new attack surface that they need to be paying attention to.

Why now? I mean, we have been doing AI… I actually taught a class in AI and decision systems in 1989. It is not like AI hasn't been around. It has become much more prevalent because the hardware is better, the math is better. We understand a lot of things differently than we did in 1989, thank god. In the five years, we have really seen the uptick in AI as a thing that is in its own right and is being used, and it is also under the covers. That is the other piece of this that I see with systems that I deal with is, there are a lot of places where the system that is not an AI system, but there is AI under the covers that is enabling certain aspects of that system to operate more efficiently, faster, et cetera, et cetera. Why is now the right time for putting together the security incident response team idea?

Lauren: Yes, really the widespread adoption that happens with things like ChatGPT and other large language models [LLM] coming out. Because it used to be you needed some type of specialized skills to really interact with the AI. You had to be in Python. You had to be in a coding language. Now all you need to do is open the browser to start asking questions. That really fueled this widespread adoption, lots of publicity into it, which really fueled the industry to start putting it into about everything. There is hardly anything I think you can buy right now that doesn't have AI labeled on it or is using AI. Then too, just with that widespread adoption, you hit a second point that I would like to make is, it ended up kind of everywhere. If you are using your cell phone, most likely you got a GPU on there right now, which is most likely running some type of AI, especially in your pictures to erase things and the background maybe you don't want to see. But it just really permeated our society. We felt like now was the time if you are ever going to jump in and start trying to backpedal and fix some of those secure development practices, or at least make people aware of it. Like I said, that is one of the main objectives.

Suzanne: Well, and I would also argue that because we have been so successful with CERT and incident response teams, from a public viewpoint, there is an expectation that any software is going to be protected, is going to be secure, is going to have fewer vulnerabilities. This is a case where that is only true if the people building those systems are aware, and they are using the right practices and keeping track of what is happening in their space. Because I am guessing that there are going to be some novel vulnerabilities that aren't even possible in a procedurally based system that all of a sudden, some attacker is going to go, *Ooh, look what I can do*, and is going to wreak havoc with some aspect of AI. The public really isn't ready for that because we have now… Since 1988, we are coming up on 40 years, '28 will be 40 years since CERT was founded. We have a whole generation that has grown up thinking that we should be able to rely on software for the most part. Obviously, there is the ransomware, and things happen along the way. This feels like this is a whole new threat space that we have got to pay attention to. If we don't make people aware of it, then they are going to be surprised, and they are going to not take precautions that they need to take to be able to protect their assets and data. Good lord, we know data is one of the things that we really need to protect so all those kinds of things.

You gave a presentation at RSA recently with our CERT director Greg Touhill. You discussed your experiences in creating this in the division, including some of the challenges and the early lessons learned. Why don't we start with the challenges. What did you see as being the challenges in getting the AISIRTs going?

Lauren: I would say the biggest one is cybersecurity folks and AI not always speaking the same language and almost being scared of each other's fields. When you went to the cybersecurity, and you are like, *I have an AI vulnerability in your software. You are an incident response team at company X. Let's handle this.* They would be like, *Oh, that is AI, that is not me*. And you are like, *But who does do that part*? So, starting to help them, you are like, *In some cases, I can even localize it for you and make it very specific. Is it in this line of your code, or it is in this configuration file in terms of the weights of your neural network or something along those lines?* That was one of those first hurdles to get over in terms of creating it. Then, I would say some of the lessons we learned being just how big the problem is. When you think about AI, I have heard software, and I will say AI software, but when you look at the supply chain, it is so much bigger. A lot of times, if you think of like your LLMs or how you interact with it, you have the software that makes it up. That software builds upon libraries, upon libraries underneath with data, somewhere clear down at the bottom. Then it is also the devices it runs on. One of the things with some of the large language models and other AI/ML [machine learning] applications are, you need it to be efficient. You need that speed. Well, in order to get that speed, we made some tradeoffs on things like GPUs where we seem to have forgotten a lot of the CPU best practices for data privacy and security reasons. They just don't scale. Then some tradeoffs were made in that area. You are also bypassing a lot of protection. Like your OS-level [operating-system-level] protections are gone because a lot of times it runs on the chip on the bare metal. We have given a lot up to meet the scaling demands that we need to. Just trying to re-get that in. We released a vulnerability note in the fall that talked about some of the GPU issues, specifically LeftoverLocals. Trail of Bits came out with a very good article and had reported it to us. We helped coordinate it across a few different vendors to start to fix those things. Libraries, the frameworks, OpenCL, OpenGL. We have a lot of ways to go.

Suzanne: Well, and if I think of something simple. I have used ChatGPT to build a meal plan that has specific parameters. I can very easily see in that very simple scenario the data that it is trained with. If it is trained with data that represents really crappy recipes, I am not going to get a very good meal plan. I will get a meal plan, but I am not going to get a very good one. That is one of the things that I am seeing in early users of AI systems is that I am pretty forgiving right now. If it doesn't give me the best recipes. If it is not giving me the best meal plan possible, I am pretty forgiving because in my mind, I am thinking, *Well, we are still building the data set and all the rest of that*. But that could just as easily because some adversary is going, *I am going to make sure that the whole United States is malnourished because I am only going to give them crappy recipes*, *when they ask for recipes and meal plans*. I am being ridiculous. You know what I am saying? That data dependency means that that vulnerability is one that we don't necessarily think of as a vulnerability. We just think it is not trained well enough. Is that something in particular you have seen, or is that something that now I am telling the whole world about that the attackers are going to go after? But I am curious as to whether that is something that you are already seeing.

Lauren: We haven't seen too much in terms of the data poisoning. There is a lot in the academic research, but we haven't seen any of those vulnerabilities come across. That may be because they are being exploited by adversaries. But we have certainly talked through some of those use cases. Meal plans being one of them, one that we have talked to is actually languages. If you look at things like a lot of the current large language models, they are heavily trained on English. If you wanted to be able to bypass, then maybe you think about some of the not-as-used languages or slang. You are going to really trip it up. We have talked about even when incidents get reported in those types of scenarios, debugging, what went wrong?

Suzanne: Right, right, right.

Lauren: It is huge, right? Because you are like, *Well, did it just hallucinate? AI is not going to get it right 100 percent of the time. Is it one of those cases? At what point does it start to be a pattern that needs further looked into? How do you look into it? Do you have to start introducing a lot of that data to improve it? Do you have to go back to the drawing board or just retrain?* There is a lot of discussion to how to even figure out if it is just, AI is not going to get it right 100 percent of the time, or if it is truly an issue. And if it is an issue, is it an issue due to training? Is it an issue due to adversary influence or interference? Like I said, the supply chain causes those questions to get a lot longer to get to the bottom of it.

Suzanne: Oh, absolutely, and supply chain risk in this area has got to be a nightmare. If you are looking for using something in an official capacity, in a business capacity, understanding those supply chain risks is going to take on a whole different scale, as you said earlier. Data centricity is something that…One of the other areas that I am doing a little bit of work in is studying software development lifecycles for AI engineering. I am sort of hyper-focused on data right now, but that seems to be one of the real differences between traditional software engineering and AI/ML engineering is this reliance on training data as one of the key elements of success for your system. Are we starting to educate the security aspects of that? Because there is the CERT. There is the incident response, but I know that you go beyond that. It seems that the education of both the cybersecurity specialists in terms of, *Hey, here's your new threat space. Welcome to our world*. And also, the AI engineers in terms of, *Hey guys, if you are dealing with this kind of data centricity, you have a new world to deal with in terms of the kinds of practices that you are dealing with*. Are we starting to see that happening as well? I am looking for like a cybersecurity certification that says, *includes AI content*. Are we there yet, or are we still too early in the evolution of AI as a discipline?

Lauren: Sure. We just passed about six months in terms of the AISIRT team [Artificial Intelligence Security Incident Response Team]. We have really been comparing to what traditional CSIRTs did, to see where are the gaps. Where is AI introducing new things? I like to say our initial kind of look is about 80 percent is a direct fit and about 20 percent needs to adapt. Data is definitely an area that needs to adapt. That is really where we focused our research and training too. We can do some of the overall training, but it is a lot easier to say, *Hey, the mechanics are the same, cybersecurity community. This looks just like a CSIRT. These pieces are the same,* but it is really where we have been focused is on those new pieces. Certainly, data is one of them. I know I have a lot of conversations with folks about SBOMs, how they need to adapt for AI.

Suzanne: SBOM is a software bill of materials, for our viewers that don't know that term. They should know that term but not all of them do yet.

Lauren: We have been talking about, is there an AI bill of materials that starts to include the software bill of materials but also data of things like model cards too, like some caveats of, this is what my model was designed to do. This is what it was trained to do, so if you start to take it outside of what it was meant to do, people know. And they know that they shouldn't trust answers.

Suzanne: A big part of this is being explicit about what it is that we can be, what are the vulnerabilities that are possible, and what are the steps we can take to make sure that those are dealt with. Then, also, as you say, things like, *this is what the model is meant to do*. As soon as I started looking at recipes—I'll go back to my non-business model—as soon as I started looking at recipes on ChatGPT, my mind went immediately, *Ooh, meal plans.* Now, as it turns out, it seems like there has been some training in that area. But maybe six months ago, if I'd have said, *Hey, give me a meal plan,* it wouldn't have been able to do it because it was outside the training of what that particular model was ready to do. So we are going to have that same learning curve, I think, in other areas where as soon as you see something that AI can do for you, you think of the next thing. You may think of a next thing that the model developers didn't think of. So understanding what you can expect, what is reasonable to expect from the large language model or other system, is part of educating people that know, *That is not a vulnerability. That is just what this was meant to do*. *It wasn't meant to do X. It was only meant to do Y.* That is good for people to think about.

Lauren: Oh, I was going to say, even the model cards that have been developed are pretty complicated. We also think through how to make that more digestible for folks or something that they even know to look at.

Suzanne: Yes. I don't even know if ChatGPT has model cards for meal planning. That would be interesting. There is a whole class of these models that essentially learn through usage. If nobody asks about meal plans, then that model is not going to gather information about it. As soon as they see an uptick in people asking for meal plans, the way they work is they will ingest more data related to that that they can find. There is this relationship between usage patterns and actual data acquisition patterns in these models that is very different from procedural software. We are not used to the software making a—I don't want to say a decision—but making a choice to ingest different aspects of data. We generally expected to just use the data that we fed it and not go out looking for more data. That is one of the things I see as being a difference in that whole data-centricity aspect of AI.

Lauren: I would say you just stumbled upon one of the key differences in the vulnerability world too. Because a lot of times I can't replicate a vulnerability for that exact reason. Because it has been retrained, and we don't even snapshot models well. I can't replicate it for you. If I report a vulnerability, one of the first things a company wants me to do is give them the proof of concept, so that they can replicate it. But now I can't do that because they have already retrained, and they don't really publish what version they are under, or they don't always update a version when they've retrained. I can't tell them the version, so they can't go back to the date. And I can't localize where it happened in the code because I interacted with it through an API and sometimes for an LLM. So yes, this is where it starts to get complicated, especially in the generative AI world.

Suzanne: Every time I do one of these podcasts… I won't sleep tonight. You give me all of these things, so many more things that can go wrong. But, on the other hand, these are tools to help [make] our lives better, work better, and make our missions work better. They have a place, but we really do need to be vigilant about what are the cybersecurity aspects of this new tool that we have in our toolbox for making things better.

Lauren: Yes, and especially in the AI world too, like giving those protections to the security researchers, so they can poke the holes and find the things the developers never thought of to make it better. We actually have some really great ones out of Carnegie Mellon. Matt Fredrikson and his team have put out some really great research. We have things like the universal and transient attacks, which what they are is basically suffix attacks. If you ask an LLM something like how to destroy humanity, it will come back and say, *Hmm, we can't do that,* right? But then, all of a sudden, they put some extra characters on the end, and it gives you a blueprint. There are a lot of large language model companies that want to make sure that doesn't happen. Actually, Matt's team did the right thing and reported it to them. They said, *Ah, it is fixed*. Then he put some different characters at the end because he has got an automated way to do that and broke it as fast, well, probably faster than they fixed it. Then that starts to get some more of those architectural… How are they architected? We can't just keep outlawing certain regular expressions because researchers like that can find them faster than they can probably put the regular expression to say, *If it matches this, don't generate an answer.* Yes, trying to get them into the cybersecurity world, think about the design choices, and help even flag for folks that more research should go into that so we come up with better solutions than just slapping more Band-Aids.

Suzanne: Yes, and so the Band-Aid approach is actually what we see early, right before we really understand some of the underlying principles, underlying frameworks, underlying architectural constructs that govern how these things work. I think we are still kind of in that discovering what some of those are for AI. Not everybody remembers when we were in that phase when we were looking at complex procedural kinds of systems, but we were there. We went through the same thing where the first set of fixes was explicit and coded to a specific set of characters or a specific set of expressions. Then you figure out how to abstract and what is the principle that is underneath this that we need to guard against, not just listing every single possible variant of that expression. I resonate with that having come from that '80s-era kind of software development. All right, so if I am chief information security officer or whatever title I have, and I am in some kind of enterprise, whether it is DoD or commercial, should I be looking to set up my own security incident response team? How big do I need to be before I need one of these as a separate entity? Is there enough guidance out there yet for me to do it myself? Do I need to come to CERT because it is still early in infancy? What do I do if I am actually concerned about this at the top level, and I want to make sure that I do the right things for my enterprise?

Lauren: First of all, I think asking the question, *Are you leveraging AI anywhere?* for whether or not you need an AI security incident response team? It is not all these…

Suzanne: Is there anywhere where the answer is going to be? If you ask me that question, I can't think of any organization that would say, *No, I am not leveraging AI. I don't allow AI in my organization.*

Lauren: You would not believe how many times I hear that like, *No, we don't have AI anywhere.* And I am like, *Well, do you use like general products that most companies would use? Are you on a Zoom right now? There is probably some AI in the background to do all the different gestures and things that are along those lines*. No, most people are using it in some capacity, but it is trying to get them to understand where. That is usually the first kind of gate that we work to answer with the teams or with different companies. Then the next one will be like, *OK, now maybe you know where you are using it. How prevalent is it, if it is a couple of places*. No, probably not a whole entire AI security incident response team but make sure you know your product security incident response or your PSIRTs, your CSIRTs. Know who to call, like know that AI is being used in that area of the company. If I see something wrong, I can reach out to them, and we can have a discussion. Same, flipside, I tell them, if you think, *Yes, we are going to integrate an LLM to make everything easier*, bringing a cybersecurity person to that conversation so they know that you are doing that, and you deployed it somewhere within your company. When something does go wrong, they have some awareness, and they can make sure they are logging things coming out of that too. We have a lot of discussions that way. If you are an AI company, you should have an incident response team with lots of AI capabilities. You should probably be red teaming that AI. I'll have a whole list of suggestions. I mean, it is really company specific, but we are always here to call too and talk through all those different use cases.

Suzanne: Are we going to be seeing, so in the CERT world, the traditional cybersecurity world, we went through a phase where it is pretty much enterprise-level CERTs. Then we actually went to national and international

CSIRTs. Do you think we are going to see the same kind of trend here where we will need some higher-level coordination of things as AI becomes more and more prevalent, or have you already decided how you are going to set that up? Is that part of the plans right now or are you kind of waiting to see how things evolve?

Lauren: Yes, we are still in that research phase where we are not 24/7 operational either, right? We are trying to figure out where the hard problems are, where we should be focused, and at the same time, try to start to introduce some training to your national CSIRTs, right, in terms of AI.

Suzanne: Oh, yes.

Lauren: Yes, so we are trying to more incorporate it into the existing teams and create separate teams. We are really separated from, like I said, for that name for awareness. So like, *Oh, I have an AI problem. I'll call them. And then I'll tell you AI is software, and it is largely not different and…*

Suzanne: Go back to your team and tell them how you are dealing with secure coding.

Lauren: Exactly.

Suzanne: OK, all right, that is fair. Over time, we may see some different things emerge depending on how this whole area evolves. One of the things, as a federally funded research and development center [FFRDC], the SEI is really about transitioning ideas into the public and our government stakeholders. You have this wonderful RSA presentation that you and Greg put together. We'll link to that. There is a video of that as well that RSA has provided. We are going to build a blog post on this. We build blog posts on almost everything that is of any importance. In addition to those kinds of resources, what are some other resources that our audience members should look for or organizations they should be following besides CERT to make sure that they are aware of what is going on in this AI vulnerability space?

Lauren: Oh, that is a good question. I would say not going outside of CERT initially to answer your question. We do post vulnerability notes, just like we do for software vulnerabilities, on AI vulnerabilities through our website, kb.cert.org. So monitoring there to see if you are running some hardware or different models that we are coordinating vulnerabilities with and disclosing. Similar to that, we are putting out some coordinated vulnerability disclosure guidance for AI/ML, so it is close. We are hoping in the course of the summer it will get out. I will say the hardest part there is, every vulnerability that comes in I think we learn something new. The team is constantly like, *Oh, but there is just this one more point I need to work on*, so I think it will be a living document for us. I don't know how it could be anything else with how fast this field is changing. It will be something that will continually get updated. Then if you want to look outside, oh, there is a lot that is good in terms of the different red teams, what other companies are doing, things that they are thinking about. There are different AI vulnerabilities that pop up in the news quite frequently as well. For me, I almost have a hard time keeping up with that volume of information.

Suzanne: I can tell you just researching AI software development, system development, lifecycles, just about the time that I think I've got the whole landscape mapped out, it is like, *Oh, and did you take a look at this that just got published?* I think that is actually one of the messages that we want to make sure people understand is, this is a very—you have said it before, but I am going to reemphasize it—this is a very fast-moving field. We are at that stage where we have kind of got hyper-acceleration of this domain. Somebody has to keep looking. You cannot look away for even a minute. If you really want to stay up to speed with this. If it is critical to your business, you are going to need to pay attention to this. It is going to be an ongoing basis. It isn't just a once a year, go to the RSA Conference. This is something you are going to have to—which, go to the RSA Conference—but this is going to be something you are going to have to pay attention to all the time.

Lauren: Yes, absolutely, and I would say look in a lot of the places you would look for cybersecurity things. So CISA [Cybersecurity & Infrastructure Security Agency] put out with, oh, I can't remember how many seals were on it. It is a lot, secure development practices for AI developers. If you are on the development side, there is a lot of guidance coming out. I would say the guidance on the user side, there is a lot of that too, depending on where you are looking, so yes, but there is something new every day.

Suzanne: You are going to be busy.

Lauren: Already busy.

Suzanne: Already busy. You have got the operational aspects of getting this stood up. What is next for you on the research front? Are there any particular areas where you are going, *Ah, we really need to look at this, and I am the one to do it*?

Lauren: Oh, narrowing down that list is pretty challenging. One is trying to come up with some secure development training tailored for AI developers, because that is the heart of a lot of the problems that we see. It is a part of a lot of where we see the adversaries attacking. Like you could come up with a really fancy data poisoning attack, but when the door is wide open over here, like, why wouldn't you just use that? Yes, so from a…

Suzanne: Good point.

Lauren: Yes, from a cybersecurity perspective, so we really want to start getting that cleaned up, and then that really gets you like secure by design, going after those principles. I would say, one of the biggest ones for us, and then the other one is just continuing to raise awareness that these vulnerabilities are cyber vulnerabilities. Trying to get everyone to take advantage of the cybersecurity practices that we have had for a long time and just evolve those to support AI, not come up with something new. In one area that we didn't talk about, harms and biases, there is a lot of debate going on in the AI community. *Are those vulnerabilities? Should they have their own process? Should they follow the existing cybersecurity ones?* Those are actually conversations that I am probably in weekly to try to figure out how we are going to tackle those. A lot of times, I will admit, we lean towards vulnerabilities because by the CERT definition, it violates an implicit or explicit security policy. That is where we lean, but we have a lot of conversations in that area. Those are some of the big ones. Then, like I said, too, when you go with awareness, just how big the supply chain is. We see everything from attacks, like I said, on the hardware to the models themselves, to how you interact with the models through things like the APIs. Trying to have people understand, because sometimes they are like, *Oh, well, it is just the model piece*, and you are like, *Oh, no, it is so much bigger*, so just bringing general awareness to that.

Suzanne: Lauren, if people have vulnerabilities or they think they have vulnerabilities in their AI system that they are using or they are developing, what should they do about it?

Lauren: We request that you report those to us at kb.cert.org. The more vulnerabilities that are reported, the more that we are able to learn where these are similar and different and how to tackle this challenge. Then that really allows us to update things like our CVE guidance for AI/ML, so we can share those best practices with all of you.

Suzanne: As I said, you are going to be busy. I do want to thank you for taking the time to talk with us about this today. To our listeners, I want to thank you for joining us today. I am hoping that we have given you some ideas about things you should be paying attention to, even if it is just to figure out where is AI in the systems that you use and that you produce. We are going to link our transcript to all the resources mentioned in the podcast, and that is our standard practice. We will not stop doing that. I do want to say that this podcast series is available in all the places you can find podcasts: Apple Podcasts, SoundCloud, Spotify, and of course, my favorite the SEI YouTube channel. As always, if you have any questions, please don't hesitate to email us at info@sei.cmu.edu. Thank you for joining.

Thanks for joining us, this episode is available where you download podcasts, including SoundCloud, TuneIn radio, and Apple Podcasts. It is also available on the SEI website at sei.cmu.edu/podcasts and the SEI's YouTube channel. This copyrighted work is made available through the Software Engineering Institute, a federally funded research and development center sponsored by the U.S. Department of Defense. For more information about the SEI and this work, please visit www.sei.cmu.edu. As always, if you have any questions, please don't hesitate to e-mail us at info@sei.cmu.edu. Thank you.