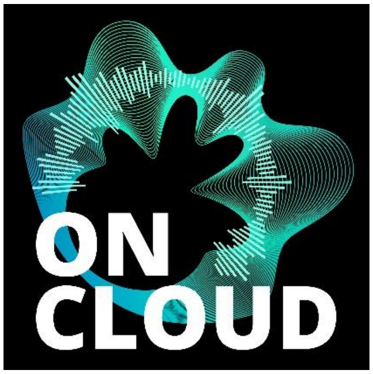
On Cloud Podcast
March 2024



The Deloitte On Cloud Podcast

Title: **Google’s Connie Fan and Deloitte’s Dr. Abdul Rahman on Generative AI and cybersecurity**

Description: In this episode, Connie Fan, product strategy and analytics lead at Google, and Deloitte’s Dr. Abdul Rahman examine the growing relationship between Generative AI and cybersecurity. They explain how Generative AI can boost cybersecurity through predictive modeling, and they discuss the need for enhanced security when training AI models—especially the LLMs used in Generative AI. Finally, they forecast the future of AI and cyber and discuss the importance of assessing AI and cyber risk.

Duration: **00:26:30**

David Linthicum:
Welcome back to the On Cloud podcast. And in this show, I am joined by Connie Fan, product strategy and analytics lead at Google, and Abdul Rahman, applied AI leader at Deloitte. How are you guys doing?

Abdul Rahman:
Great.

Connie Fan:
Doing great, David. Thanks so much for asking.

David Linthicum:
Yeah. It’s great, it’s great to have you here on the podcast, because we’re going to talk about security, and AI, and really kind of where this whole thing is going. If I get one question out there from my clients, and from the press. and people I talk to at Deloitte and outside of Deloitte, it’s, “How do we secure these systems that we’re building fast and furious in the cloud with all the explosion of Generative AI that’s going on right now?” In fact, I noticed that the cloud conferences have kind of turned to Generative AI conferences, even though they run on cloud.

So, it’s going to be a world, where probably for the next several years, we’re fast and furiously moving toward AI, and our ability to secure it is going to be paramount to what we’re looking to do. So, let’s get into who you guys are first. So, Connie, tell us about you and what you do at Google.

Connie Fan:
Yeah, sure thing. Hi, folks, I’m Connie Fan. I am the product manager of SecPaLM, and a security specialist here at Google.

David Linthicum:
And Abdul?

Abdul Rahman:
I’m Abdul Rahman. I am an AVP within Deloitte. I lead the development of AI models for cybersecurity use cases.

David Linthicum:

Awesome. And is that development of AI models that we're protecting, or development of AI models to participate in protection of data, systems, things like that?

Abdul Rahman:
It's both. It can encompass a wide variety of things, in addition to operationalization of models that can both represent different parts of the computer network operations, or CNO triangle, as in building models to help support offense or defense, or really trying to understand a little bit more about the infrastructure, whether that's on-prem, cloud, or hybrid.

David Linthicum:
So, kind of what I'm summarizing here, Connie, your role is really kind of understanding how AI exists in the domain of Google, and how you guys are working on it. And Abdul's role is really to take all different technologies into practice and implement them to the best value that the clients have to offer. So, I love this kind of thing here, where we're talking to someone who's dealing with the technology itself, and someone who's technology agnostic, has to understand how all the stuff kind of fits together, and what works and doesn't work, and how it works into a professional service chain. So, Connie, going to you first describe the security, AI security strategy at Google. And what are you guys working on these days?

Connie Fan:
Yeah, I would say Google, writ large, is working on a bunch, in terms of how to secure AI development, like you were mentioning. Lots of exciting stuff going on there. The small sliver that I focus on, right now, is when you want to apply AI in the security domain, of course, there are a lot of general purpose models out there, and a lot of open source frameworks that are amazing for so many use cases. But security is such a specialized domain, the question for my team is, what sort of work do we need to do, above and beyond the more general-purpose work, to really give security professionals and experts, like Abdul, who are making this tech come to life for end customers that much more useful.

And, so, for us, it started from a selfish place, if I'm being honest, when Google teams first started looking at LLMs, and saying, "Should we put this in our own products? Would it be useful to our users?" About a year ago now, when we were experimenting with the general purpose models, we were encouraged enough to want to pursue this, but we weren't getting high quality enough responses. And, so, we needed to fine-tune a model to be more security-focused.

And then it kind of occurred to us, if we need to do this for ourselves, anyone who wants to build a LM feature, or integrate it into their own work internal workflows, so long as that problem itself is focused on security, they will then have to fine-tune a model likely as well. With now the inclusion of the Mandiant team, they've got such good data in the virus total team. And we really think that the various groups across Alphabet now have some of the best security telemetry out there. And when data quality really dictates model quality, we thought we were in a really unique position to do some of that fine-tuning, how to tie up and offer a more tailored solution to security folks.

David Linthicum:
Yeah, one of the things I did last weekend was watch probably 10, 15 YouTube videos on Google security's journey. And, so, it was is interesting to me just kind of catching up on that. And the Google conference was a few months back, really trying to get a good understanding of how all these things are kind of fitting together and working, where the advancements are being made. And it's incredible, the fact we've kind of gone from, say, 35 miles per hour, in the AI space.

We always knew is going to have value. It's been around for a long period of time, to probably 80, 90 miles an hour now in the industry, and everybody is moving fairly quickly. And, really, kind of the ability to think of all the pragmatic stuff, security governance, data management, all these sorts of things, is something that I would rather not be an afterthought. So, Abdul, going to you, in looking at this technology and what they're building, how do you evaluate it, and look for it to be applied in a certain space? And what are you looking for, specifically, as you piece these things together to create solutions?

Abdul Rahman:
Well, we like to ground a lot of our model development in some practical frameworks. One aspect of our development focuses on building models to support, detect, and respond workflows within the Deloitte MXTR platform. Within that sort of sphere of what we're doing, we could, for example, look at the MITRE ATT&CK® model, where the book ends, which are really the tactics at the far ends of the MITRE ATT&CK, really correspond to network telemetry, whereas the inside portion tactics correspond to really endpoint agent data collection.

And, so, a lot of our sort of evolution of models, in terms of understanding, rides on a few, what we'll call "holy grail principles." Among them, is being able to connect, and corroborate and correlate network data with endpoint data, so that we can actually see the whole flow of the conversation, not just from what's occurring at the endpoint, or different endpoints, but also have a conversation as occurring for different kinds of cyber phenomena, like, for example, lateral movement. So, we've built a series of models that are built in ways where we can reuse components. For larger models, we kind of build them together like Legos.

So, in the case, where we're trying to detect, for example, zero-day threat, we connect that with a lateral movement model. And we also train this on data, that is typically both audit log, identity and access management, along with network traffic data, in a way where we can actually see, not just very, very specific, in other words, looking for needles among needles in haystacks, but really tried to understand what are the possibilities for adversarial workflow can occur. And that includes insider threat.

So, some of the more complicated problems where organizations face either AAPT-level, or nation-state-level threats, concerns detecting signal that is attributed by low and slow activity. In other words, adversaries blend their actions in the noise. And, so, a lot of the data that Google has collected, or other threat intelligence providers, gives good understanding of selectors. But the sense-making aspect really has to do with being able to connect it from an operations standpoint, and actually what you're trying to measure.

David Linthicum:
So, let's expand on a few things that you mentioned here. Number one: use of AI models. Some of the cloud pros on there may know a lot about cloud, may not know as much about AI. Could you define AI models for us? And also walk us through a detect and response scenario. This is Abdul.

Abdul Rahman:
Okay. So, there's a lot of conflation around AI. But I would say an AI model is a way build a prediction, a method code, a codification of a technique, or business rule or mathematical equation on top of data. And, so, when you look at something similar or simple as linear regression, which is computing the best fit line across a set of data, where there may be sort of a linear type of behavior, the line would represent potentially the next prediction of points, or probably a range of where points could exist.

So, the way that we like to think of AI, in the case of cyber is, really we're performing correlation, we're performing forecasting, we're doing some degree of prediction, for very, very specific use cases. Now, there are some AI models that are built a little bit differently than others. So, there's, at the core of it, looking at AI from a standpoint of maybe logic. There's deductive reasoning, and then there are things like abductive reasoning. In the case of abductive reasoning, you're looking at evaluating whether something is true, but it may have some, potentially, aspects around it, or some assumptions, which may not necessarily be true all the time.

And in cyber we deal largely with stochastic capabilities. So, if we were to describe AI from a code perspective, it's fundamentally stochastic code, where stochastic really represents a probability or probability density of a particular activity occurring. Now, in the case of detect and respond, there's really two aspects of it. The response part could be instrumented potentially using a SOAR capability (security orchestration, automation remediation tool) of which there are plenty, and they have AI and machine learning embedded in them.

On the detect part, we're really looking for signal. So, within the actual data, if the data is labeled, or attributes, and when we talk about labels and attributes, we really are talking about something equivalent of the headers on Microsoft Excel columns. So, those can be considered labels. And what we do with these labels, or attributes, is formulate a path toward looking at how we can construct them or organize them in a way, combine them in a way, concatenate them in a way, to formulate things like interesting aspects that we want to measure, or phenomena. Or in data science, we would call those "features."

So, if we understand the data, and as Connie mentioned, data is of a certain quality, we're able to build the stochastic representations of data to predict things based on features, which are fundamentally phenomena we want to measure. So, all these fancy words are to say is, "Hey, the data that we collect, and the subject matter expertise of cyber analysts, and operators, all comes together," to say, "Hey, when we are looking at traffic or network traffic, and we see certain behavior at certain times of day, at certain ports, where we're seeing an abundance of maybe syntax on TCP traffic, then that typically represents this kind of intrusion set."

So, we would build a model to basically model that effect. And in some cases, very, very specific models, where we know the phenomena, that's what's called "known indicators of compromise," that lead to sort of known outcomes. Then there's a class of models where we have sort of known IOCs. And we don't know what the output is. It could be a variant of malware. And then there's capabilities where we don't know what the IOCs are. We know something's happening. And there's sort of an unknown response. And, so, that's really the known knowns, known unknowns and unknown unknowns.

And typically towards the side of the spectrum, where you have more known unknowns, and unknown unknowns, that's really where behavior-based, or these stochastic techniques and models shine, because we're not typically using rules for behaviors of threats that already exist. So, that's really kind of how you would want to bring it together from a standpoint of AI, and really, maybe potentially looking for something interesting in the detect and respond world.

David Linthicum:
Outstanding. Connie, I'd love to hear your take and how you're dealing with data management. And at Google, and obviously, data is kind of core to AI-based security and your ability to not only secure these models, but also secure the data that's encapsulated within these models. And also the fact that we've seen that data can change. Data can go from innocuous data to something that really needs to deal with compliance and needs to be protected. So, how are you focusing on data management and threat management at Google, and what are some of the core components and technologies you're building right now to take that whole thing to the next level?

Connie Fan:
Yeah, it's a great question. I think, first and foremost, the most important thing for us is a principle, not a technology. We do not train customer data. If you give us data, whether it be in a prompt or any other data that we have, we will never use it in training. So, you will never see a pop up in response of a model to someone else, or even to yourself. I think in terms of technology for data management, one thing I'm really, really, really excited about is the use of extensions, particularly in the LLM space. And, so, Google calls it "extensions." Folks like Open AI call it "plugins."

Across the Generative AI space right now, it really is just the core concept, though, that LLMs are not just generating text, but they are able to call an API. And, so, to one of Abdul's previous points of, there's a lot of great threat intelligence out there. But to operationalize it is an entirely different question. To be able to securely call the APIs of your intel provider, to be able to extract from that API call what the IOCs, or indicators of compromise, like Abdul is mentioning for a certain threat actor. And then have the LLM stick that into whatever syntax for whatever investigative query is relevant to the tool that you're using, to be able to do your day job, is like a very quick turn that was previously inaccessible.

The crux of the question is that the best way to secure the data is to not touch it at all. And when we need to call it, we're calling it from the already secure API. And then the ability to do that is not just purely for the standpoint of securing the data, but really to be able to access fresh and timely data, and sensitive data, to be able to make that very useful in the outputs for various users.

David Linthicum:
Awesome. So, Connie, what has changed since we've started going with standard – standard security to protect data, tech processes, protect behavior, things like that, into security and AI, what are the fundamental changes that occurred within not only the technology that you're producing at Google, but the industry in general?

Connie Fan:
Yeah, it's a good question. I am of the mindset that, in fact, many of the threats that folks had to be concerned about previously, are still the threats that they have to be concerned about now. I think in terms of what's really changed is just folks' awareness of how important it is to protect your data, particularly with some of the recent leaks, various individuals within companies maybe not being fully aware of what happened with data that they submitted to these models, and then being unpleasantly surprised when their data showed up in the response to someone else. It really has just raised some level of awareness and some level of vigilance out there. But I think in terms of the core problems that people need to be aware of, fundamentally have not changed.

David Linthicum:
Yeah, I think I think that's a great response. Because I think that then what the message should be, people who are listening to this podcast, is ultimately security is still important. We're just changing the aspects and tactics. Abdul did a great job explaining what those things are. And Connie did a good job explaining on what weapons or what mechanisms you can take to make these things happen, or to provide better security. So, Abdul, I'm going to go to you. So, what's the progress that's being made? What are the fundamental changes that are occurring now and will occur in the next three or four years in terms of how AI-based security, AI is related to security, is going to morph and change?

Abdul Rahman:
That's a great question. So, there are movements to operationalize AI down narrow paths. And then there are broader ones, which I think will lead to a lot of fruit. And, some of these broader paths are roads that are not as well traveled. And among them is really understanding of attack surface management. We've been investing a lot in trying to develop attack surface cartography through reinforcement learning. We've published about seven papers in that area. We're finishing up a book. And essentially what we've learned by trying to understand attack surface is, really, attack surface is not a static thing. In other words, adversaries constantly examine through reconnaissance and other measurement, other measures, the current state of a target, and how it could potentially change.

So, some of these interesting aspects that we're looking at is, what are the dynamics of attack surface? How do they change relative to maybe IOCs? How do they change relative to updates to security or patches? And how do they change relative to tools being deployed, for example, access through APIs, to small language models, large language models, the security focus large language models that Google supplies and others. I think part of what we're going to see in the future is gravity moving toward some of these more relevant foundational pieces of cybersecurity among them. How is our attack surface evolving in an organization?

Other portions have to do with simple questions that some CISOs have trouble answering, which is related to attack surface, but more is related to risk. Right now, today, if you were to ask five CISOs the definition of risk, you probably get at least five different answers. And that's because risk is largely subjective. And, so, what we envision with AI is getting better bounds or constraints, or the ability to set conditions on what risk means in the context of CISOs being able to answer requests from their shareholders, or board members or other members in the C-suite around what, what is the risk of data exfiltration? What is the risk of us being hacked?

What is the risk – what is our security posture in light of this new strain of this malware that's wreaking havoc, of which I think a lot of the Google capabilities and tools will be able to, be able for us to move faster. So, there is speed for response. And then there is the ability to process signal in a way where we can do things like attribute, also to do things like really acquire intent. And, so, those two, attribution and intention, are among the holy grails in cyber, which I think round out a lot of areas where AI will kind of head us into.

David Linthicum:
So, Connie, walking the halls of Google probably a dozen or so times in my career, I always got involved with these great conversations there in terms of where the technology is going and really kind of asking probing questions. And, so, what are the conversations that you're having right now with your colleagues, without revealing any confidential information, in terms of where this stuff is moving to, and where the work needs to be done to make this stuff better?

The fact of the matter is we're going to have dynamic threat vectors that are there, and the ability to attack vectors that are there, and the ability to adjust to them. And also the ability to adjust to the fact that hackers are out there weaponizing AI to come get your stuff, and in many instances, leveraging the same technology and the same tools.

Connie Fan:
Yeah, it's a great question. I would say, all of that. And then even beyond that, not necessarily what folks on the outside will do, or the non-users will do to attack the users. But what will the users hurt themselves with? I think the thing that we're very much concerned about right now, and actively pondering, is how to reduce hallucinations. So, that when we give responses to customers, they can run with it, and more novice folks won't just take incorrect information and run with it. That's not something that we want. And that's a form of risk in and of itself. That is true for the security LLM. That's true for all of the LLMs, and all of AI, generally.

I would say for us, another thing I've done a lot and said about stochastic modeling, is when it is great, it is so great. I think you also hear people accuse LLMs of being stochastic parrots, and they just generate the next word for you, and it sounds close enough. But you don't know if it's right or not. Being able to wrap the model in other more core systems, engineering to support it, to guide it towards the correct responses, is a large area where we're focusing now, and we are very excited about.

David Linthicum:
Yeah, I think there's a bunch of ways in which you can move in this direction. And right now, we have so many opportunities in front of us because we have so many problems that I think need to be solved in the, in the AI security space. And really kind of fundamentally and creative thinking needs to occur. And we really need to kind of fail fast, and having mechanisms and ways and approaches to do this in better and better ways. So, Abdul, where can we find more about you on the web? Where do you normally post? What do you want to tell the listeners of how to reach out to you?

Abdul Rahman:
Well, we operate in a large team or a cluster, which is a group of development teams, for building cyber AI models. We've published a lot of papers in the past 18 months. We've done a lot of conferences. So, I think generally, a lot of our team has presented, and it has their papers and listings within the archive, Arxiv dot-organization. We also have a lot of presentations and publications within IEEE Transactions, along with Google Scholar. So, I think just being able to look at the different domains where we build things, we've made a lot of contributions to using AI for ransomware detection, for a zero-day threat, for lateral movement. We've also built analytic capabilities for attack surface management.

And we are collaborating on the next-generation MDDR platform to be able to acquire these signals at near real-time and be able to process these detections much faster. Leveraging, of course, awesome capabilities and datasets from Google and others. And that platform is called "Cybersphere." So, we're actively working both internally and externally, from an eminent standpoint, along with being able to work with our core partners to move the ball forward.

David Linthicum:
Wow. That is a lot of content out there. And I've written from any of those publications themselves, they get into a lot of detail very quickly in terms of how you do something, not just what it is. So, Connie, same question to you. Where can we find more information about you and what you guys are doing at Google?

Connie Fan:
Yeah. Before jumping into me, I just want to say I've read some of Abdul's papers myself, and they are amazing. We so value Deloitte's partnership here at Google. For myself, personally, you can find me on LinkedIn. But I really want to give a plug to the broader team. Google released Gemini very recently. And we are so, so, so proud of that offering. Yesterday as well, Duet for Chronicle, duet for SecOps is now GA. And, so, that is the set of features that really make Generative AI available to the fingertips of defenders. That that is now out there for everyone to use, and everyone to kick the tires on.

So, the main thing I would ask listeners to do is just get hands on is kick around the tires of our products and let us know what you think. That would be super helpful feedback for us. And we're really keen to hear more about what users want.

David Linthicum:
Absolutely. I drank from the Gemini firehose this weekend, and learned a lot of multimodal AI modeling, and some of the cool technologies that's coming from that. And it's kind of tough keeping up trying to make sure that we're not only keeping up with the latest and greatest in some of the AI stuff, but also the management and monitoring and security we're talking about here, and how well that's changing around the use of these technologies. Absolutely amazing. It's absolutely the greatest time to be in this industry, in my opinion.

If you enjoyed this podcast, make sure to like us, rate us and subscribe. You can also check out our past episodes, at deloitte.com/us/cloud-podcast. Now to share a personal note with all of you, although you will be hearing my voice in upcoming episodes of the On Cloud podcast, I am officially stepping away from the microphone. The show will continue to bring the inside stories and unique perspectives you've all come to expect from the On Cloud. Thank you for being such an indicial part of this journey, I encourage you to keep tuning in. Best of luck with your cloud projects and stay safe.

Operator:
This podcast is produced by Deloitte. The views and opinions expressed by podcast speakers and guests are solely their own and do not reflect the opinions of Deloitte. This podcast provides general information only and is not intended to constitute advice or services of any kind. For additional information about Deloitte, go to Deloitte.com/about.

This publication contains general information only and Deloitte is not, by means of this publication, rendering accounting, business, financial, investment, legal, tax, or other professional advice or services. This publication is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action that may affect your business. Before making any decision or taking any action that may affect your business, you should consult a qualified professional advisor.

Deloitte shall not be responsible for any loss sustained by any person who relies on this publication.

Visit the On Cloud library
www.deloitte.com/us/cloud-podcast

About Deloitte

As used in this podcast, “Deloitte” means Deloitte Consulting LLP, a subsidiary of Deloitte LLP. Please see www.deloitte.com/us/about for a detailed description of our legal structure. Certain services may not be available to attest clients under the rules and regulations of public accounting.

Please see www.deloitte.com/about to learn more about our global network of member firms.