Barefoot Innovation Podcast with Behnaz Kibria, Director of Cloud Public Policy, Google

***Note that transcripts may sometimes contain errors and that transcript timing notations do not match the posted podcast**

Jo Ann Barefoot: We have a really exciting show today. My guest is Behnaz Kibria, and she is the director of Cloud Public Policy at Google. Behnaz, welcome. I'm so excited to have you on the show.

Behnaz Kibria: Oh, thank you, Jo Ann. I'm excited to be on.

Jo Ann Barefoot: We're going to talk in a few minutes about the fact that we together are putting out a second paper on model risk management with AIR and Google working together on it. And it's been fantastic to be able to work with you and your team on these cutting edge topics. And I will link in the show notes to the first paper as well and the second paper. But I wanted to start out by asking you to talk about yourself, what's your background, and then tell us about your corner of Google.

Behnaz Kibria: Sure. Thank you, thanks so much, thanks for having me, and very, very happy to be working with you again on this paper. So I am the head of a team called Product Initiatives at Google, which is part of the government affairs and public policy team. And we cover Google Cloud issues in particular, and as the name suggests, product initiatives, we work very closely with the Google Cloud product teams to help them navigate policy issues. And as you can imagine, as generative AI has become an increasingly important part of our product suite, the policy issues around generative AI are also becoming increasingly important. So we are spending much more time on these issues ourselves.

Jo Ann Barefoot: Right. And what's your own background?

Behnaz Kibria: So I have been at Google now seven and a half, almost eight years. I came out of... I'm a trade lawyer by training. I used to work in the administration in the Obama administration doing trade work, and then before that on the Hill doing trade work, and before that in private practice doing trade work. So tech really was sort of a bit of a pivot for me, but it's been a really interesting eight years or so in tech. As you can imagine, a lot has changed over that period.

Jo Ann Barefoot: That is for sure. So I want to start by talking about everyone's favorite subject, generative AI. And in the first paper, we didn't talk much about AI. It was sort of the basics of model risk management, MRM, but obviously generative AI is changing everything around us. And so I think it'd be great if you would explain what it is, how it's different, and how it's being used in financial services so far.

Behnaz Kibria: Well, you're asking the critical foundational questions, and as you know, this is the sort of the foundation of the paper too. So the key to understanding what

generative AI is is that term generative, right? So generative AI is the type of artificial intelligence that can create new content or data. So instead of just analyzing or acting on existing information, it can generate things like texts, like poems or articles or summaries or even code. It can generate images, pictures, artwork, photorealistic scenes. It can generate audio. It can generate video. And what really sets it apart is that generative AI models learn by being trained on massive amounts of data, and these models identify patterns and structures in that data, and then use the knowledge that they acquire through training to generate new, sometimes similar content.

And you mentioned the first paper that we did last year and now the paper that we're working on in that period, one of the things that's really changed in the world is really the bursting onto the stage of generative AI sort of evolving from earlier traditional models.

And when you look at, we did a study back in July 2023, a benchmarking study and asked organizations about the impact they felt GenAI would have. And 82% of the organizations responding that were considering or using GenAI believed that it would either significantly change or transform their industry. So that gives you a sense of how the industry, how industries are looking at the technology.

In terms of the benefits of generative AI, it creates the possibility of enhancing individual productivity, strengthening security operations and advancing data-driven decision making, so that all of those capabilities really come together in terms of financial services to create a number of compelling use cases. One you can imagine is enhancing customer service. So I'm sure everybody's experienced AI-powered chatbots and other sort of personalized recommendation engines and things.

There's the possibility of streamlining operations, automating tasks, streamlining workflows, and we've got a number of customers of Google Cloud who are already using generative AI for those purposes like Deutsche Bank. There's the possibility of improving risk management. So for example, fraud detection and risk assessments. There's the possibility of using AI to boost investment research, so analyzing market data and generating investment ideas. And then there's also the possibility of creating personalized content such as marketing materials and financial reports. Those are just five sort of general categories in which you can see real potential in the financial services industry.

Jo Ann Barefoot: Yeah, absolutely huge. I think our listeners have heard me say this before, but I've never seen a tech trend that was big and hit everyone at the same time before ever. If we think about the other big ones, they kind of gradually came into our lives over years, and this one within a few months, that little few month period, two years ago or so, everybody was figuring out what this thing was and playing with it and thinking about whether it was going to change what they're doing, so [inaudible 00:07:02].

Behnaz Kibria: Absolutely. I mean, if you compare January of last year to January of this year, it's really just a really significant change.

Jo Ann Barefoot: Yeah, absolutely. So you work in public policy as do we here at AIR, and talk about the regulatory environment around AI, but specifically generative AI. Is it fit for purpose for what we're going to need in terms of the legal frameworks and the regulatory frameworks and particularly in financial regulation?

Behnaz Kibria: That's a great question, Jo Ann, and we actually started getting that question even before this burst of energy and excitement around generative AI. As you know, because we work together a lot, when this issue first started bubbling up in the context of more traditional AI models, we started getting questions both from the regulators on the one hand and from our customers on the other saying, with all of this new and emerging technology at that point still sort of traditional AI, do we need new laws and regulations? And we took that question on board, and the first paper that we were, as you know, really was an attempt to answer that question with respect to a subset of traditional AI use cases, what we called at that time risk use cases, the use of traditional AI technologies for fraud detection, anti-money laundering, et cetera.

And the question that we asked was with that set of emerging AI use cases as a sort of a basis, did we feel that we needed new laws and regulations? And the answer that we came to in that context was no. And the reason is because one of the great characteristics of the laws and regulations in the financial services industry is that most are principles-based and technology-neutral. And what that means is that they are capable of being adapted over time as technology changes. Every new turn in technology doesn't require that we throw out the tools and start again.

And as we looked at the use cases in that first paper, we studied in particular SR 7-11, which is the model risk management framework in the United States, and which was issued in the wake of the global financial crisis long before AIML models were in the mainstream use. And while that was updated in 2021 to issue more specific guidance on AI, the principles themselves in that kind of framework as we looked at them really did seem sufficiently high level to provide the kind of meaningful guidance that we needed in the assessment of risk even as technology advanced. And so the reason for the second paper now is really to look at that same question again and ask, with the continued evolution of technology, is there a need, does the same conclusion hold?

Jo Ann Barefoot: It might be useful to our listeners to talk just a little bit about, even though many of them I know already know, a little bit about what we mean by a model in this context, and what are those basic principles of standards and expectations for a sound model or model risk management process?

Behnaz Kibria: It's a great question. I think that the way that models are defined, especially in the US system, it's actually quite broad. This is not the legal explanation for it,

but I think the way that it's been applied, it's really anything that's got inputs and outputs could potentially be a model. And there are certain categories that model risk management frameworks look at, things like model validation, assessment of a model's accuracy, reliability, and limitations. That's one category of issues that model risk frameworks look at. Governance and control, so roles and responsibilities for model development, implementation, and monitoring, that's another category issues that model risk management frameworks tend to look at. And then risk mitigation, which is identifying and managing potential risks such as bias, data quality issues, and misuse, that's another category. And so the general framework of model risk management that we have in the US and replicated elsewhere as well is sort of passing new technologies through that frame in order to assess and mitigate risk.

Jo Ann Barefoot: Tell me if you agree with this, but I think one of the issues that people are kind of waking up to is that not too long ago, model risk management and model validation was sort of the province of a expert group of data science and tech people who really understood how a model is built and what it's doing and what it's supposed to do. And as these tools get better and stronger, it's becoming more and more possible for non-expert laymen to use them. And that causes a new set of challenges in terms of things like governance or things like who's keeping this model up to date and so on, and more risk probably that the person using the model may not be fully alert to what could go wrong with it. And so I think it's extremely timely to be doing what you're doing.

Behnaz Kibria: Yeah, I agree with that entirely, Jo Ann, and I think you're headed towards one of the issues that we highlight in the paper as being unique or different with respect to the world of models that we're dealing with now, which is that you also increasingly have more collaborations between technology companies like ourselves and financial services firms developing these models. And so you've got sort of very similar to what you said, you've got many different actors potentially in the supply chain or the use chain and with differing understandings or differing sets of expertise, and so you've got to have a framework that covers all of that.

Jo Ann Barefoot: Note to the editor, I'm making a departure here that our little script says there's a question for me. Do you want to ask it of me or do you want me to jump in on it?

Behnaz Kibria: I'm happy to ask you. Because this is a collaboration, I thought it'd be interesting to sort of turn the table on you.

Jo Ann Barefoot: Okay. Just give it a three-second pause and go ahead.

Behnaz Kibria: Okay. So Jo Ann, I wanted to ask you because we've been collaborating in this space now for a number of years and AIR has been doing even other work in the AI and GenAI space. What are your takeaways and drivers for further work? How

important do you think it is more broadly to share MRM frameworks and principles based on the conversations that you've been having?

Jo Ann Barefoot: Yeah, thanks for asking. I cannot think of anything more important than the whole ecosystem that we're working in leaning forward and preparing for the changes that are coming with this technology. And it'll be like anything else, it sometimes seems really fast and sometimes it seems slower than people predict and so on, but there's so much work ahead to figure out how best to do this for all the players in the ecosystem.

We think that generative AI in particular is really going to bring an entirely new paradigm of risk management and financial regulation because people will have more opportunity to see more information and analyze it easily. I mean, the biggest... If you think about the core function that risk managers are performing is they're trying to figure out what is happening that could be causing problems. And it's hard to do that, information is hard to get to, and it's hard to analyze it efficiently.

Generative AI changes all of that combined with the greater digitization of data, which has been exploding, doubling frequently for a long, long time. And one of the things that we think is going to develop is a lot of use of agentive AI, AI agents that are helping people do what they have to do, including find out what they have to find out, so the ability to query with using an AI, a complex set of data and get insights and advice out of that without being stuck in sort of static reporting and that type of thing.

But the challenges are just massive, and that's including for the regulators. And I'd like to throw it back to you again, when you think about the challenge here for policymakers and regulators, what are the most... and we do have a lot who listen to the show by the way, what are the most important things for them to be thinking about?

Behnaz Kibria: Oh, that's a great question. And I think one of the things that I think is really important is to remember that we can have new technology, we can have very, very advanced technologies that really challenge old assumptions and old ways of doing things, but that doesn't necessarily mean that the tools that we have can't be useful. That's really just a core premise of this paper is that even though the technology is different and really just revolutionary, the ways we've been looking at structuring our processes, our thought processes around risk are still pretty valid. And what's really important is to break down the issues into component parts and then to assess what's truly unique about the technology that requires a different look, a different way of addressing things? And as we think about this new paper, that's really what we've tried to do is to take the frame of model risk management that's been developed, impose it upon the technology, and then identify the very unique aspects that require sort of a special consideration.

One example is just in the context of safety and soundness. In the old days when you had traditional models, you could just unpack how the model worked. You could just say if X, then Y, if Z, then Q, and that's how the model was built. And if the inputs and outputs kind of match what you would expect based on how the model is built, then that's sort of sufficient evidence of safety and soundness.

With the evolution of technology with generative AI, that kind of unpacking of the model to understand its guts becomes a harder task. These models are very, very complex. And so that kind of assessment of safety and soundness is no longer valid. Then you have to move to what are other ways of assessing safety and soundness against this new technology.

And one of the core proposals of the paper, for example, is that you really have to look more at things like testing, output testing. You have to look at techniques that are utilized in the process of model development like grounding to really assess safety and soundness. So those are new concepts, but the frame against which they're put, the way that we look at evaluating safety and soundness, the structures, the way of thinking about them is still a really useful way. It's just that the best practices, of course, will evolve as technology evolves.

Jo Ann Barefoot: When we think about the challenges facing regulators in particular, and our listeners know, I always say I think that regulators have the hardest job. They really don't have leeway to try things out that aren't going to work and that type of thing other than sometimes in their labs and sandboxes.

I think one of the challenges that's going to really be hard for them is putting out guidance that is both clear, but can be made agile too. They don't like to put something out before it can be clear. I mean, sometimes they'll put out very broad guidance that requires subjective judgment, but they're trying to go for as much clarity as they can generally speaking. But this area is changing so fast, and I don't know what the solution to this is, but I think they need to be thinking about it, the ability to put out guidance on what they expect in, say, MRM for generative AI.

And then as they learn more and as the field evolves further, put out something else without creating too much chaos and too much sort of version control, because what if the regulated entity was following the guidance of the time, but now the guidance has changed? That's fine if you have pretty stable situations, but when you have a very dynamic situation, it all becomes harder. And I think that's going to tend to make the regulators take longer to reach clearer guidance, and I think that in turn will cause a lot of problems in practice. So I think there's a tremendous challenge there. But I'll say at the same time... Go ahead. Go ahead.

Behnaz Kibria: No, I was just going to agree with you. I think that's right, but I want to hear what you were suggesting.

Jo Ann Barefoot: I think the point that I made about transforming the risk management paradigm, I think that goes doubly for the regulatory and supervisory processes themselves, and the fact that there will be so much more transparency in these complex financial systems and transactions. We're going to be dealing with the challenges in crime and fraud that are being caused by AI, which is definitely a huge issue, but we're also going to have massively enhanced ability to see them. It's going to get harder to hide them in the system if the regulators are equipped themselves with AI detection tools that can find patterns that we humans don't have enough days in the hours in the day to look for.

And I'll throw in one more idea that we've been thinking about, which is if consumers become equipped with their own AIs to help manage their financial lives, their own AI agents, that could transform the way we think about consumer protection, if there were an ability for the consumer to easily see through a predatory practice or a product that was a poor fit for their situation or a risk to their privacy or something like that, you could imagine that some of the things that agencies do now to protect consumers might be less needed. They maybe should focus on regulating the AIs and making sure that they've got the consumer's best interest in heart. So I think it's a whole brave new world of possibilities as well as threats for the regulators and their missions.

Behnaz Kibria: That's such a good point. And actually you asked me earlier how is AI being used in the financial services industry? My favorite use case is actually one that it echoes what you were just saying, which is the risk management uses of AI. We spend so much time talking about managing the risk of AI, and yet AI has such a useful role to play in helping financial institutions manage risks as well. So I'm very glad you made that point.

Jo Ann Barefoot: Yeah, good. So let's go back to the financial institutions. You talked about some of the emerging use cases that are exciting. What are the things that you and the folks working on the paper have been coming up with in terms of what financial institutions should be doing?

Behnaz Kibria: Well, as we look at risk management, we sort of bucketed into four issues, four areas in which we think there's a variety of best practices that financial institutions should be considering and that regulators should be maybe acknowledging as evidence of good activity or best practices, developing robust governance and ongoing controls, things such as integrating sociotechnical considerations and stringent data governance. Because again, as we talked about one of the really unique things about AI technology is just the importance of data. Good data in means good outputs as well.

Another areas is supporting reliable model development, and their techniques such as grounding techniques and outcome-based evaluations, as I said earlier, take on real heightened importance. Strengthening model validation and oversight, and here in particular the really important aspect of human in the

loop systems for critical decision-making to make sure that you do have the necessary extent and level of human oversight over these processes.

And then the third we've also talked about before is it is very likely, especially with complex models that the relationship, there's going to be some sort of a third party involved in the mix, some sort of technology provider is likely going to be providing some support along the way. And so as we look at risk management and the relationships between all of the participants in the process, there's got to be a clear delineation of roles and responsibilities and a sense of sort of shared fate and shared responsibility as well.

Jo Ann Barefoot: Right. Is there anything we haven't talked about that you'd like to add?

Behnaz Kibria: Well, just maybe another question for you, if you don't mind. Following your recent work on modernizing regulatory agencies, what steps do you think regulatory agencies can take to keep up with the kinds of advancements we've talked about?

Jo Ann Barefoot: Some of it is common sense. Our listeners know I'm a former regulator myself, I was Deputy Controller at the Currency, and at the financial regulatory agencies, all of them have very strong cultures and deep expertise that they're very proud of, and kind of a sort of dominant professional way of looking at types of things and so on. And I think one of the most important things they need to do is very explicitly and proactively bring technology right into the center of those value systems and ways of thinking.

And it's not easy to do, but it's probably getting easier because of demographic change as more young people are moving up the ranks at these agencies that just sort of automatically makes it a little bit easier. But the future of finance is technology and therefore the future of financial regulation is technology. And this technological tool is the single most important one, operating as part of a larger thing in terms of creating this ability to leap forward.

So I think the agencies need to be training on it, learning about it, hiring people who understand it into leadership roles, experimenting with it, putting together experimental lab-type situations where they try using some new things and see what happens, learn along with the industry, go to all the events, listen to the podcast shows, just do everything possible to bring this into the DNA of these regulatory bodies and get everyone to immediately think of tech.

Working in this field as I have been doing for so long, I never have more than a few days go by where I don't hear somebody talking about something in the regulatory realm, whether it's a private sector or public sector person, where to me, the issue obviously has a tech potential solution, and it's not in the toolkit of the people who are working on it. They don't mention it.

And usually, again, it's these kinds of things, it's just the ability to learn more faster and understand what it may bring faster. That's what regulators are doing all day long. Everything they're doing is really at heart understanding risks and non-compliance and so on in the system and trying to find it.

So I just think that there just needs to be an all-out campaign. And it also needs to include, and this doesn't get discussed enough in my view, it needs to include that the regulators have to update their own tech architecture. They are still trying to operate, as many banks have been trying to do, grafting new tools on top of an outdated tech stack that you get to something that's better than nothing, but they all have to move to cloud computing in one form or another, and they all have to get all their data into... and all the data they're using, internal and external data, structured and unstructured into their ability to work with it, and really overhaul their own tech.

And that's a project that's going to take some resources to do that. To me, that's all the more reason to get started sooner rather than later. And they all are. I mean, we have some US agencies that are very sophisticated with AI, and we have some that are just starting on the cloud journey, and a lot in between. But nothing is more important for regulators, it seems to me, than getting on top of both the risks that are emerging with technology, but also the ability of technology to, A, solve for those risks, and B, solve for problems that we've always had in the past and have not been able to solve, timeless problems like how difficult it is for the current system to detect money laundering, for example.

The numbers on that are just shocking as to how much money laundering goes undetected in the system, the ability to find patterns of potential bias, the ability to find emerging risks to the safety and soundness of an institution, or the systemic risk of a sector or a sub-sector in finance, the ability to make sure that the consumer protection rules are functioning the way they're supposed to be.

We didn't used to have the ability to do any better than to use the tools that we had, and therefore, a great deal of our regulatory philosophy is based on looking at the quality of the inputs and the quality of the controls as you've been saying. So the regulators expect to see a robust compliance management system and the three lines, if it's a bank, the three lines of defense and all of that.

And again, that was the best we could do before, but now we're getting to the point where we're going to be able to see outcomes. We have consumer protection laws. How are they doing it, protecting people? Might we be able to do better? We should be really rethinking the whole thing in light of these powerful new tools that have been put into our hands. So I'm very optimistic and excited about it. Regulators all over the world are thinking about these questions, and the momentum is all toward figuring it out, but it's a mountain to climb in terms of exactly what to do.

Behnaz Kibria: Yeah, no, I couldn't agree with you more. And I do think there's a paradigm shift probably that's needed, not in terms of necessarily the frameworks, but the way of thinking, right? And it's been changing already, but even if you think about the period of time we've been working together or that I've been working on cloud issues, the nature of the conversations really changed. I mean, just even as a cloud provider, we used to be often outside of conversations. In financial services, those conversations often were between the regulators and the regulated entities, the banks, the financial institutions, and we were sort of an extraneous to that conversation. That's really been changing, and I think that's positive. I think there is a three-way conversation that must happen as technology becomes more and more integrated into banks and financial institutions.

And then just the nature of the conversation too, where it's much more fluid and collaborative rather than sort of being like a tennis game where one side hits the ball, then the other side hits the ball, then the other side hits the ball. It's a much more fluid conversation between three people or three entities rather than that. So anyway, all to say, I agree with you entirely, and I think there's a lot of good stuff that is on the horizon.

Jo Ann Barefoot: Right. So we will put in the show notes at regulationinnovation.org the new paper, the previous MRM paper. We'll put in a couple papers that AIR has just put out, including one we did a podcast on recently on what we've entitled The Financial Regulators' Odyssey, written by our colleague, Nick Cook, and some of the other tools that we've been working on here at AIR and anything else that you have to share from Google. And again, that'll be regulationinnovation.org.

Behnaz, I can't thank you enough for joining me today. Behnaz Kibria, it's just fantastic to have you, and I hope the paper will have a big impact.

Behnaz Kibria: Oh, thank you. It's just been a real pleasure working with you and being on the podcast.

Jo Ann Barefoot: Wonderful.