

Barefoot Innovation Podcast: Delicia Reynolds Hand, Senior Director, Digital Marketplace, Consumer Reports, Dr. Adrienne Heinrich, Head, AI Center of Excellence at the Union Bank of the Philippines (UnionBank) and an AI Advisor at Aboitiz Data Innovation (ADI), and Dr. Henry Balani, Global Head of Industry & Regulatory Affairs, Encompass Corporation

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

Jo Ann Barefoot: I am so excited about today's show. I've been looking forward to it for months because we have three guests today who submitted papers to us in AIR's call for papers on generative AI and the financial sector and especially the consumer financial sector. And we have had three fantastic papers. Needless to say, we will have them in the show notes as well for you, but we're going to spend some time talking today to the three authors.

Our guests are Delicia Reynolds Hand with Consumer Reports, Dr. Henry Balani from Encompass Corporation, and Dr. Adrienne Heinrich from the UnionBank of the Philippines. And I want to welcome you all.

I actually would like to start by asking each of you to just tell us a few moments about yourselves and then the title or subject of your paper, and then we'll go in more depth from there. Delicia, let me start with you.

Delicia Hand: Sure. Hi, Jo Ann. Thank you so much for having me here. My name is Delicia Reynolds Hand and I currently serve as the Senior Director of Digital Marketplace at Consumer Reports. I've spent my career at the intersection of consumer protection, financial technology, and public policy. So from being at the CFPB leading consumer initiatives to now being at Consumer Reports where I'm spearheading research and other initiatives with a focus on making financial systems fairer, smarter and more inclusive. So at CR, that work includes a focus on AI accountability, digital finance, research. As you know, we had started to test and evaluate popular consumer apps, FinTech apps, payment apps, digital banking, etc. So that is my background. Did you want me to say anything about the paper?

Jo Ann Barefoot: Just what's the title of it or the one line subject of it?

Delicia Hand: Yeah, sure. So the big picture of this paper is about how consumer financial protection needs a major upgrade for the AI era. So not the formal title, but it really looks at what I try to look at as the paradigms of 20th century consumer protection, specifically looking both at regulatory and statutory foundational

laws and how right now it has produced a system that reacts when things go wrong and what are some opportunities for the 21st century. So what are some 21st century paradigms driven by technology for the future of consumer protection?

Jo Ann Barefoot: I'm going to come back to you in a moment and we'll go deeper there. I will also say you've been on the show before and we will link to that in the show notes as well. Henry, can you introduce yourself?

Dr. Henry Balan...: Yes, absolutely, Jo Ann. Well, first of all, thank you for having me, I truly appreciate it. And very pleased that you found my paper worthy of a presentation and a podcast. So absolutely happy, delighted to share that with you. So I'm the Global Head of Regulatory Affairs, Industry and Regulatory Affairs that Encompass, and what Encompass does basically is to support the financial institutions to help address financial crime. And the focus is around understanding some of the challenges that banks have, especially when they're looking at onboarding customers and so forth, and providing solutions. And we do that through something called a corporate digital identity. And one of the things that I was looking at when I was looking at putting the paper together was really trying to understand how do we create more efficient processes around helping financial institutions. Now we work with the large global banks and so on, but this certainly applies to many of the other financial institutions out there in terms of making the whole onboarding process.

In other words, identification. Not necessarily from individual perspective, but from a corporate perspective. Companies need to be onboarded too. Companies need to buy financial services from banks and so on. And the question really becomes how do we help the banks do that? From a technology perspective. So as a regulatory technology firm, that's what we help them with, being able to help automate those processes to get them to where they need to get to. But I think what I was looking at as I was writing the paper, was trying to get a sense of what does it mean from an open standards perspective? Because ultimately the belief and the understanding is that you got to have standards, established standards in place to be able to help address that. And that's what the title of my paper is, is really the role of corporate digital identity standards to address financial crime. So that's what I dig into and happy to chat a little bit more about it as we go through this podcast.

Jo Ann Barefoot: Absolutely. Really looking forward to it. And Adrienne, I want to thank you for joining us. Delicia and I are on the East Coast of the U.S. and Henry is in the central U.S. in Chicago, and I know you're in Manila and at least what, probably 13, 14 hours different from our time zones here. So we appreciate you joining us in the evening, your time. Go ahead and introduce yourself.

Dr. Adrienne He...: Okay, thanks, Jo Ann. Really excited to be here and also happy to do that in my time zone. I'm Adrienne. I am heading the AI Center of Excellence for UnionBank of the Philippines. I'm also AI advisor for Aboitiz Data Innovation. I'm really

passionate about AI innovation. I have been busy with AI use cases for a very long time, for almost two decades. And what I really enjoy about the podcast and the topic that we could submit on is the other perspective. I'm actually an AI developer, but now we have proposed a GenAI governance framework. So that's actually really interesting for me to bring the two together, innovation and regulation, and that's also what the paper is about, about navigating GenAI ethics and our proposal from a technology standpoint, what should be in place in the different stages of GenAI development life cycle.

Jo Ann Barefoot: Yeah, I'm really excited that you are bringing a developer's perspective to this discussion, which is partly about regulatory and standards in general. So let's jump into the discussion. And again, Delicia, let's start with you. You and I have had a lot of time over the years to talk about how to protect consumers. As you said, you were at the US Consumer Financial Protection Bureau when it was founded, and I'm very struck by your line of thinking that we are in need of a new paradigm for how to protect consumers, leveraging technology better. Talk to us about your high-level thinking on how we move from a reactive to a proactive approach.

Delicia Hand: Sure. So fundamentally what we've seen, and I talk about this in the paper, that consumer protection has been reactive for a long time. So what I mean by that is we wait for harm to occur before acting. And even if we're not waiting, we fail to take the steps to really set up consumer-facing industries and businesses to know how to protect consumers. And so we've been continually in this cycle of a big crisis happens, millions of consumers are impacted, then you have regulatory or other action, a new law is passed. And so effectively a proactive model would flip this and also would use technology to anticipate risks and intervene early. And so in the paper I ask us to think about imagining and frankly we don't even have to do it. We don't have to imagine anymore because as technology does, given how mainstream GenAI is inaccessible, we are seeing on a daily basis the introduction of AI-powered everything.

But imagine AI-powered financial assistance that monitor at the individual level, spending, flag predatory practices, can intervene and negotiate on behalf of consumers in real time at the individual level instead of consumers navigating complex disclosures. They don't read them anyway. So really thinking about the consumer experience, but again, thinking about how moving from this individual-level, agentic AI potentially can change consumer business interaction as well. So in a world frankly, in the U.S. anyway, deregulations, I think, these kinds of tools can become even more critical and give consumers the means to safeguard their own financial well-being without relying solely on traditional oversight institutions and that might actually level the playing field and build a healthier marketplace.

Jo Ann Barefoot: So this topic has fascinated me for years, but we didn't use to have good technology for going after this idea of agent AI helpers for consumers. You've used the term financial guardians, which I love. Talk to us more about how you

would see this working. What would be the business model if you can envision that? One thing I worry about is whether if consumers were using generative AI to guide their choices, make good choices, prevent harm and so on, that they might be using a tool that had an ulterior motive, a secondary agenda. How are you thinking about how this model might evolve?

Delicia Hand:

Yeah, so there are challenges and opportunities that we would need to optimize and regulate. And so maybe to the challenge side, I think what I'm hearing is you're implicating a few things. One primarily is trust and privacy. So a financial guardian which could travel with the consumer, helping them manage moment-to-moment, day-to-day and really revolutionize and hopefully aid the consumer in achieving financial security sounds great, but they come with challenges. So these systems would need for example, to access financial data so that implicates trust and privacy because we'd need to ensure that security and privacy standards are essential. So perhaps zero knowledge proofs could allow AI to provide these insights without seeing raw data. We would have to also address bias and fairness. AI, we already know from mortgage lending and other spaces, can reinforce existing biases.

So if not properly monitored, this would get out of hand and effectively replicate current disparities. So you would need regular audits and algorithmic fairness standards to be part of the job or the tool. And maybe to your earlier point, the business model matters. So the AI should not just fit into an existing consumer-facing tool. In my view, this should be the only job of the AI. It should be truly independent, it should empower and not be embedded in services that are being offered to consumers such that the consumer can retain full agency over the tool and have autonomy over its decision-making and participate in the decision. So there's a lot of business model questions which are implicated here. On the opportunity side aside, I think there are very specific jobs that these guardians could be doing. They could, for example, help consumers in real time know when they are likely to overdraft or help to prevent debt traps with real-time intervention.

They could also enable consumer control financial profile. Why do you need a credit report that is static, maybe inaccurate, et cetera? This could be a new form of credit report that the consumer can actually access in real time and might reduce reliance on bias credits scoring. But the first job that this AI tool should have would be to be loyal to the consumer themselves. So from a business model perspective, we would have to think about things like is it free? Freemium access to core protection? Are there ethical partnerships? So is this governmental or non-governmental so that it's truly independent. We would really have to think about the business structure. Is it consumer-owned like a credit union or does it have nonprofit funding? But keeping the model transparent and fully aligned with consumer wellbeing would be the first start.

Jo Ann Barefoot:

I think this is so important and exciting. I've worked in consumer financial protection for decades and the design that we have for protecting consumers,

there are some hard rules and prohibitions, but the main model over the years has been disclosure, and then there's been a lot of effort to supplement that with consumer education, financial literacy education. I think both of those are important, but we also know that they don't get the job done, that people still struggle to manage their financial lives and understand financial products, which are just inherently complicated. I had someone say to me the other day that, "Finance is too complicated for AI." And my reaction was, "I think it's too complicated for humans." Maybe the AIs can fill in some of what we're missing.

Before we turn to Henry, you also have touched on how you're thinking about the role of blockchain technology going forward in consumer protection and the decentralization of so-called trust. Talk to us about how you're thinking on that.

Delicia Hand:

Yeah, I'm a student of blockchain, so I know less here. But really when I talked a little bit about this, it was really thinking about the role of banking institutions and how they might shift with technologies like blockchain. So I don't think traditional banks will disappear, but their role will shift and potentially instead of being gatekeepers, maybe they may serve more as trust anchors in a decentralized system. So if you think about blockchain and smart contracts automating traditional banking tasks like credit verification and transaction processing, this would force banks to compete on service quality rather than infrastructure control.

And so tokens, new coins and all of that aside, really the rise of DeFi, decentralized finance, could mean that consumers would have more financial autonomy, but each technology creates new risks. And so potentially there is a world where AI could also help to bridge this gap by simplifying [inaudible 00:18:50] entirely complex DeFi protocols to help consumers benefit from blockchain technologies. So I think especially in the fraud and scam space, these could be leveraged to help consumers avoid scams and actually make financial literacy more active based in the transaction and so more meaningful to consumers. So it's that AI-driven transparency leveraging blockchain is what I was driving at in the paper.

Jo Ann Barefoot:

There's so much more in your paper and I hope we'll have some time to circle back to some of it before the end, but I do want to turn to Henry, and that was a great segue because, Henry, your paper is focused on a very different use case for generative AI, which is looking at the financial crime space and the anti-money laundering challenges and specifically the Know Your Customer rules or KYC rules. Talk to us about what's the big picture on how you're thinking about that?

Dr. Henry Balan...:

Yeah, so it's interesting. I've been in the financial services technology industry probably about a long time and it's always interesting to see the evolution of technology in terms of the different tools and so forth to help support that industry. When you come to AI and specifically GenAI, I think you're seeing really a significant step change, as my colleagues in the UK to call it as well, because of

the potential impact that it has, not only on the... mostly on the efficiency and the automation or the processes that we are looking at here. Because at the end of the day when you look at banks specifically that needs to comply with regulations, it can get fairly complex in terms of the needs and so forth. Regulations also change over time, and as we see different sanctions come in, we see different types of emphasis in terms of how governments and regulators are using these types of tools to be able to correct bad behavior by airing countries for example.

And so it becomes a real complex challenge for banks to be able to keep up with not only the regulatory side, because ultimately the whole job of KYC and customer onboarding is to, number one, address the required regulations so you don't onboard a terrorist or some bad person out there. And number two, to make sure we assess the risk of that particular entity. Now the entity could either be an individual as a consumer, as Delicia was talking about, but also looking at other corporations because our corporations after all need bank accounts too and they need access to financial services and so on. And so the level of complexity, as you can imagine when you look at a corporation specifically is magnified significantly in some ways. And so it gets much more complex in terms of the process by which a bank needs to onboard one of these large corporations, and this is where GenAI comes in. And here's what I mean by that.

When you look at the number of documents that a bank needs to review in terms of being able to assess not only the reputational reliability but also whether they are on some kind of a list of some sort, a do-not-do-business list, sanctions list or whatever the case may be, the issue becomes how do I actually pull together the right information, the right data points, the right attributes to be able to help us understand what the true risk assessment can be. And so there's the obvious things like the name, the address, the location of the company, the type of company it is, but also some transactional history, credit worthiness, ratings, et cetera, et cetera, and you get all this data coming in from different sources. If it's a large public corporation, sure you can get it from, for example, exchange listings, or if a UK-based company, you can get it from something called Companies House.

These are different registrars of business information that can be used to pull in that relevant data. Now if I'm working with a global organization, the type of data sets coming in from different jurisdictions can be different, can vary and can be fairly complex. And number two, it's not as structured as you want it to necessarily be to be able to make that proper risk assessment, that call. When you start using AI, generative AI specifically, you're starting to see real good benefits here. What if I could take a document, well say, a PDF of some sort that was maybe in articles of incorporation of some sort, a trust document in terms of looking at the ownership information of a particular corporation. How do I pull all that data that's relevant to be able to make that relationship and

understand that? Because one of the challenges that banks face is to follow the money, so to speak, and in some ways is follow the control.

We call it ultimate beneficial ownership. And follow the control in terms of being able to identify who the ultimate beneficial owner of that corporation may be across multiple jurisdictions and across the sharing the divisions, how they share information with each other and looking at the personal significant control. It gets really, really complicated really, really quickly.

If I can standardize some of the data elements that's required. So I talked about some of the types of standards that are required in terms of the types of attributes that you look for, but not only that, starting to use GenAI specifically to pull the data from those documents and say, "All right, this particular articles of incorporation meets the criteria. I need to be able to assess the risk of that particular corporation, but these are missing fields," for example. GenAI is really, really good about that. So I've got a set of rules, for example, that I can prompt the tool to be able to identify, "These are what I'm looking at."

If this particular data set is missing. So say supposing the address is missing from the actual entity itself, physical address, but there are certain jurisdictions out there that don't rely on physical addresses, they rely on PO boxes, as opposed to a natural physical address of a company. GenAI can be smart enough, can be trained to say, "All right, based on that, that's not necessarily something that's required." So I don't need to go back to the corporation and say, "Well, I need yet more information." So GenAI is not only pulling all this information in, it's like being intelligent, but processing this documentation. That's why we call it intelligent document processing, to be able pull that information together, structure what's required, and then be able to identify what's missing and then actually work with perhaps the corporations that are looking at being onboarded by the bank themselves to request that additional information.

So you get this kind of back and forth and this is all automated. Imagine the efficiency around being able to do something like that. Again, I've been in this industry a long time and there are these large corporations that have full teams of 50... and I kid you not, 50 full-time employee staff. All they do is coordinate with banks to be able to provide the information that's required for them to be able to successfully onboard their organizations. It's ridiculous amount of staff that's required because of complexities, because of the unstructured nature of the documents and because of the challenges that they have in terms of going back and forth and so on. Now we are seeing generative AI as being a real viable solution. Yeah, granted, you've got some of these types of biases, issues and so forth. I'm sure we can address that, we can talk a little bit more through that.

But the overall process, the big picture that we are talking about here really is to be able to automate that process by using these intelligent tools to be able to make sure it's much more efficient. At the end of the day, it's a rapid turnaround and being able to get on board these customers because banks are competitive,

certainly. If one bank doesn't get my financial product done in X number of days, I'm going to go to the next bank. So it's competitive. So you want to be able to be as efficient but as accurate as possible because ultimately you are complying with regulations. That's why they're there. It's just a requirement and that's a reality that banks face. And GenAI, I think it's going a significant step towards being able to help address some of these challenges.

Jo Ann Barefoot: Our regular listeners know that I spent two years at Harvard as a senior fellow researching how technology is changing consumer finance and how to regulate it, and one of the things that came up in that research is that complying with the anti-money laundering rules is the most expensive compliance thing that banks do, and we know that the results are disappointing, to put it mildly. We know that we don't catch most of the financial crime. So if we could get these step changes in efficiency that you're talking about, it not only could contain costs, but it could actually make headway in finding these terrible crimes, which we should [inaudible 00:28:20] ourselves are terrible crimes that are profitable because the funds can be laundered.

Dr. Henry Balan...: Yeah, absolutely. If I could add to that, actually, you bring up a very valid point, because at the end of the day we get so caught up in the minutiae, the 'forest from the trees' analogy that we forget why we are doing this. We are going after financial crime, we're trying to catch the bad guy. Regulations are there to help support us, not help hinder us, and that's unfortunately the world that we've gone to these days.

Jo Ann Barefoot: Yeah, I couldn't agree more.

Dr. Henry Balan...: Yeah.

Jo Ann Barefoot: You have talked about the need in this to develop what you've called open corporate digital identity standards. Talk a little bit more about what's needed and how might we get there from here.

Dr. Henry Balan...: Yeah. So let's start with the basic fundamentals. When you're looking at onboarding a corporation, there's certain things that you need. You need the name, address, location, if it's a US company, an SIC code, some kind of identifier of the company and so forth, to be able to assess that company. Again, we're trying to assess the risk of the company because that's essentially what we do. Assess the risk from the standpoint of do we need to comply with regulations from a sanctions point of view or some of those areas or are we trying to see what the credit worthiness of that customer is? So there's a couple of different use cases that we are looking at here. So the idea of actually establishing these common standards is really key because once you start getting across multiple jurisdictions, the type of data sets and the type of information that's required can differ.

I just gave a very simple example earlier about an address. In one country, for example, you just don't need a physical address, and another company you need a physical address, et cetera. Then you've got things like the naming conventions of the owners, for example. Then you have another one in terms of percentage of ownership control.

There's a global wide organization called FATF, the Financial Action Task Force. And what they're trying to do is provide some standards and guidance in terms of how do we actually identify ultimate beneficial owners across those areas. And so what we are trying to do is create these standards, and the reason we're trying to create these standards is because you want the interoperability between jurisdictions, number one, and between systems. So technology, we've got something called API. Some [inaudible 00:30:56] listeners understand what these are. But if I can establish standards in terms of being able to not only cross boundaries but also cross systems, that would make it a huge benefit. As well as recognizing you've got countries that have data protection standards in terms of the EU for example, GDPR is a very common regulation that protects consumers as well as corporations in terms of the data that they share.

And so what we are looking at is if we can establish these standards and have an agreed-to standard across these different jurisdictions in terms of why we are doing what we are doing, then we are starting to get there in terms of being able to actually share data across. So fundamentally open standards will help us drive greater efficiency, automation and so forth to comply with these because then we even agree to standard. We all agree, "Okay, if the address format is in certain way from a certain country, that's okay because of these reasons." And so we've all agreed to those open standards out there. And that's really what I was trying to drive in my paper itself, talking about the needs for these types of open standards that we can all agree to, that we can share that helps ultimately, as you point out, it's the most expensive in terms of KYC onboarding for a bank. How do we make it more effective at the end of the day?

Jo Ann Barefoot:

I have so many more questions for you and I hope we'll have time to come back. One of the things you've said that caught my attention was that we're trying to combat these crimes with the equivalent of a dial-up modem era of technology. We just have to modernize it. But let me turn to you, Adrienne. Again, you're not coming from the perspective of a lawyer or a government affairs person, but rather from the tech side, talking about the need for and the principles that should be driving governance of generative AI. And I know you maybe can pick up some of the issues that Delicia and Henry briefly touched on in terms of the things we need to worry about, privacy, bias. How are we going to be sure that our systems don't do more harm than good? So at a high level, lay out for us what you're envisioning.

Dr. Adrienne He...:

Yeah. Let me first talk a bit about what we know on global AI governance, what the landscape looks like. It is quite fragmented and it's regionally very diverse. It's characterized by different approaches across different nations and

organizations where each is trying to balance innovation with ethical considerations. So for example, the EU leads with binding regulations like the AI Act. They emphasize risk-based tiers and strict prohibitions on high-risk applications. Then in the US we have a combination of federal deregulation with state-level ethical frameworks like the Colorado AI Act. And this year we also saw that Trump prioritizes rapid deployment of AI with less emphasis on ethical guardrails. So that's already a bit different than what's happening in the EU. Then in China, we've got centralized state control via strict generative AI rules with their interim measures of 2023, and they have very security-focused legislation where they're balancing innovation with ideological alignment.

Then we also have voluntary ethics-driven frameworks from Singapore, like with their model AI governance framework by IMDA, and the FEAT principles by the MAS. FEAT stands for Fairness, Ethics, Accountability, and Transparency. And they also have tools like AI Verify to promote transparency and interoperability. And then we've got some other emerging economies, countries like India and Brazil, and they're drafting their hybrid models, blending somewhere between the EU's set risk assessment and then sector-specific rules. So we can see here that the international collaboration remains quite limited, and we've got competing standards like EU being more precautionary and US innovation, innovation-driven. And there are forums like the G7 Hiroshima process and UN AI Advisory Body that aim to harmonize those ethical principles while the frontier AI developers push cross-border technical standards. So this landscape reflects a struggle to balance innovation, ethics, and geopolitical power dynamics, and these are general AI frameworks.

And while these general AI frameworks provide broad guidelines, and I believe a dedicated governance framework for generative AI is essential due to GenAI's unique characteristics and associated risks. And I think that then leads to your question that you asked me, Jo Ann, a unique characteristic of GenAI is that it creates new content, which makes it particularly useful, but also risky, and its creative capacity, sets it apart from the analytical and predictive nature of traditional AI for which current AI governance frameworks are typically designed. So while general AI frameworks provide a foundation for responsible AI development, a dedicated GenAI governance framework is really essential to address those unique challenges and ethical considerations that are posed by its content generation capabilities such as misinformation, defects, or copyright issues. And that's why a one-size-fits-all approach is insufficient. That's why we propose the GenAI ethics framework in our paper.

Jo Ann Barefoot: Yeah. So you've proposed a six stage lifecycle for responsible GenAI development. I don't know if we have time to talk about all six stages, but what are the key things that need to be done uniquely for generative AI?

Dr. Adrienne He...: Yeah, so ethical principles are crucial throughout the entire GenAI lifecycle, but their application varies at the different stages, and we intended to give actionable guidance so that GenAI developers and organizations are empowered

to translate the ethical principles into very practical strategies, which should ensure responsible development and implementation of GenAI. Now I can try to summarize some of the key ethical considerations, maybe just for some of the stages in view of time.

The first one is the problem definition stage where I think here we need to clearly define the purpose and the intended use of the GenAI system so that it aligns with ethical values. So we also need to consider if the problem could be solved through non-AI means. And here, unlike traditional AI, GenAI's creative capabilities may be actually prone to be overused, and this principle then forces us to consider whether simpler non-AI solutions would be more ethical and effective that they also avoid unnecessary risks.

I also think that in this stage it is very important to involve relevant stakeholders like researchers, policymakers, ethicists, and maybe also the public in very early discussions. Since GenAI's potential social impact like misinformation or job displacement, it really necessitates a wider discourse early on. And this will then also ensure that the diverse perspectives will shape the GenAI use cases.

The second stage is data acquisition and pre-processing. I think here we should source data in a very responsible and ethical way, and here we can also address potential biases. So we need to ensure that the data sources are diverse, that they're representative for the real world scenario, and not just take what is available. This means that maybe data augmentation needs to be applied to create more various and more variations in the data to make it more representative. Because if you train AI on biased data, you're likely to have a biased AI model, right?

What I think is also very important and characteristic for GenAI in this stage is data protection and privacy. So we have to implement robust data anonymization and privacy protection measures. And this is more emphasized with GenAI versus traditional AI development because PII data or other sensitive data may resurface in the generated output of the GenAI systems. Yeah. So we have some other stages like model development and training where we are trying to mitigate risk and include more explainability. After that comes model evaluation and refinement where we also want to ensure that we're using explainability and interpretability to evaluate model outputs and check whether they align with intended ethical values and the purpose that we defined in the initial problem definition first stage. And this is in contrast also to traditional AI evaluation, I would say, where it's typically done with a focus on performance metrics like accuracy, sensitivity, specificity, and not so much with a lens on bias and fairness.

Then comes the deployment stage. That's the fifth stage where I think we need to think also of accountability and responsibility, how we deploy GenAI systems and that it may involve a more controlled rollout instead of releasing a full product immediately. So their constant feedback and control loops along with

human oversight are necessary to ensure the quality and safety of generated content before and as it is released to the public.

The last stage is monitoring and maintenance. Here also, I think we need to establish mechanisms for users to report harmful or misleading content generated by those GenAI systems so that we can correct for it. And we also need to realize that the ways in which GenAI can be misused will evolve and maintenance must then include adapting safeguards to address new threats and vulnerabilities. And for this, we also suggest to put committees or approval boards in place so that the monitoring results of AI models can be found there will be reported to those committees. [inaudible 00:42:30].

Jo Ann Barefoot: Finish your thought. Go ahead.

Dr. Adrienne He...: Oh, I just wanted to say that by integrating those different ethical considerations for each stage of that GenAI development lifecycle, we believe that the risk can be proactively managed.

Jo Ann Barefoot: Yeah. As you said, what makes generative AI unique among AI is that it's creating, it's generating something and that raises... I think one of the first concerns that the public raises all the time is the issue of hallucinations. Oh, Henry, you look like you've got something to say on that. Go ahead.

Dr. Henry Balan...: Hallucinations are red herring.

Jo Ann Barefoot: Ah.

Dr. Henry Balan...: And here's why, right? Hallucinations of course, is a technical issue. We understand that. Part of it also is the idea that what you put in is what you get out, so you've got to be better in terms of the prompts and so forth. The real issue I think is bias, and I think Delicia and all of us have raised that. That's really where the concern is. Sure, we can actually train the systems to identify the hallucinations. We got to have better practices in terms of recognizing that you just don't take the output for what it is. How many lawyers and how many... so many examples of that where it generates cases that just exist and it's shame on you for not checking those references, for example. But it's getting better. Of course it's getting better. It's not hallucinations that we just need to be concerned with, it's bias coming out of the system, especially as a consumer. I think that's really where the challenge is.

Sorry, Jo Ann, whenever I hear hallucinations, it clicks in me and says, "That's not it."

Dr. Adrienne He...: Yeah. There are indeed different risks that those gen AI systems introduce and it's important that we are aware of what are those different risks. So also for as you say, bias for example or toxicity, then we need to try to involve that in our

testing procedures that we actually try to let it produce that on desired content so that we are sure that it can work according to our desired scenario, let's say.

Dr. Henry Balan...: Yeah, I'm a finance professor as well, part-time at a major university here in Chicago, and I tell my students, "Yeah, sure, go ahead and use GenAI, that's fine, in terms of being able to understand concepts and so on." But if you're going to use GenAI to solve the expected value of a sudden, "Calculate net present value," anything like that, you're going to run into trouble because it doesn't do well in terms of handling the math. So recognize that. And that's hallucination, that's not bias.

Jo Ann Barefoot: We only have a few minutes left. I could go on for hours with you all, but do you have questions? Delicia, do you want to add anything or have a question for the other speakers that you think we should address while we have a few minutes?

Delicia Hand: Gosh, I have a lot of thought. On the hallucination piece, I think we're seeing with frankly some of the newer tools that there are now ways to address hallucination and get... frankly, you can run loops and use researcher AI, for example, to really minimize the probability of hallucination, for example. And to Adrienne's point, I would fundamentally agree, I don't think we want to live in a world where we hand over all the jobs to AI. We really have to think about where is this, where the opportunities to really leverage to maximize the outcomes for consumers, not hand over all of the functions. Because even the most sophisticated tools have limitations, and you demonstrate that all the time.

And then to Adrienne's point about... I think both your points about standards and, frankly, global standards. I think we're in a moment where we're seeing frankly, the necessity and this current moment may demand that we move there more quickly. And I would wonder from an implementation pathway, because I think that's a real challenge, with you have these differing but equally complex, in my view, systems. What would you see as the pathway to a set of global standards given various nation states moving to look more inwardly rather than global cooperation being the norm right now? That is the current question I've got, which is how do we get global cooperation regarding adaptation of some global standards for managing these tools?

Dr. Henry Balan...: And perhaps if I may, if I could add to that, Delicia, you make a valid point. I spend a lot of my time thinking about exactly this issue because ultimately when you go back to financial crime, financial crime is global. It's not local. Of course you have local financial crime, I get that, but the big money is actually the global jurisdiction where between the US and Canada, between all these cross jurisdictions where there's huge amounts of illegal money being moved back and forth. And so that's part of the challenge that we are going to end up. How do the various jurisdictions around the world establish at least a common set of open standards that we can agree to? And that's, I think, what the key is here. We are chipping away as a global community, certainly sometimes two steps forward, one step back.

But the reality is that as we start to see these types of huge financial crime being committed and the establishment of the various papers out there, the scandals, the Panama Papers, all the different types of scandals that we have out there, these things have to be open. And that's the reality of it. Open standards can help us drive towards that.

Dr. Adrienne He...: Yeah. I also feel we need some kind of a global treaty for ethical AI. It's just so rapidly evolving and it's very complex and it demands therefore a global ethical framework. Now how we can get there, I'm also not really sure. I want that we can have a dialogue between the different stakeholders, so between governments, industry leaders, ethicists, so that we can identify what are now those universal ethical principles that we all want to adhere to, and for lack of a better word, have a code of conduct for gen AI development. It seems it is kind of hard to get there to cross the bridge because of all the different priorities and agendas, right?

Dr. Henry Balan...: Yeah.

Dr. Adrienne He...: But then in the past, the world managed with the smallpox, they eradicated that during the Cold War and the US and Soviet Union, they came together to work on that together. So apparently it is possible.

Dr. Henry Balan...: Yeah, no, absolutely. I think the challenge is the ethical standards I think is harder than the commercial standards in some ways because when you look at commercial standards between different countries, we can align on interests, self-interest, or both parties in both the buyer and the seller, whatever the commercial transaction is, we'll get to a standard because at the end of the day, if I find that because the standards are not able to trade appropriately across borders from a commercial standpoint, we'll try and address that. So I think there's more hope on that side. The ethical standards, I think that's really hard because by definition, our ethical standards are different, and when you start getting across jurisdictions, even more complex.

Jo Ann Barefoot: I wish we had more time. Our goal today in large part was to attract people to come and read these papers, and I'm sure our listeners can see why we chose three papers. You have such distinctly different problems that you're solving for in your papers, and yet these common themes on what the challenges are and what the principles are that need to be emphasized. The papers are all available at AIR's website, which is regulationinnovation.org. Let me ask where people can get information about your own organizations. Henry, starting with you.

Dr. Henry Balan...: Yeah. So you can actually Google Encompass Corporation and it's encompasscorporation.com, and that gives you a lot of information about us as an organization, our focus around solving the issue around corporate digital identity. So you type in 'corporate digital identity Encompass,' you'll find us there. Or just type my name. I've got a whole bunch of papers I've written on this, and you'll find me up on the World Wide Web or LinkedIn.

Jo Ann Barefoot: Excellent. Adrienne.

Dr. Adrienne He...: For me, also, you can connect with me on LinkedIn. I'm really enjoying any interesting thoughts. Also, thought-provoking comments, reach out and I'd like to have a discussion. Or if you want to collaborate, I'd be really happy to. My organization is UnionBank of the Philippines, so you can also go on the website to understand more about the bank itself if you're interested. Or Aboitiz Data Innovation, which is another business unit in the same conglomerate of Aboitiz.

Jo Ann Barefoot: Great. And Delicia?

Delicia Hand: Yeah, I would also say LinkedIn. Actually yesterday, my colleague Ben Moskowitz, who runs our innovation lab and who's my partner in justice at Consumer Reports, put out a call for folks to come talk to us. We're working on some consumer-facing solutions driven by agentic AI. So please look us up, myself, Delicia Reynolds Hand or Ben Moskowitz, and then check out Consumer Reports. Consumerreports.org is our website.

Jo Ann Barefoot: I cannot thank you all enough. I have gotten so many ideas from talking with you, and this is just the beginning of this conversation. Obviously we're just at the beginning of this amazing journey, so thank you all for joining us. I can't tell you how much I appreciate it.

Dr. Henry Balan...: Well, thank you, Jo Ann. Really appreciate the time as well.

Dr. Adrienne He...: Yeah, thanks a lot. It was great discussing this also with the other authors of the other papers. That was a really nice session. [