

Ken:

AI is of course a pretty hot topic and there's one national leader who's also a psychiatrist. And John Torous is a friend of the family and was an engineer in college, became a psychiatrist, did Informatics Fellowship and is really kind of one of a kind person on AI. When I wrote NAMI's first book with a lot of people's help, John Torous was in that book talking about mental health apps. Were they safe? Were they a good idea? How to think about mental health apps. A fun trivia question is our book was published about six weeks before ChatGPT was released. And so the entire first book that NAMI ever wrote was before the advent of AI. And so it's one of the last books in 600 years of human history that you know is written just by human beings, which is kind of a fun fact, but it does speak to how quickly this is all moving.

That book was published, our book was published in September of '22, and ChatGPT followed about six weeks later. Let's go to the next slide, please. I want you to remember you're not alone. If you need to chat with somebody, this is not a suicide crisis prevention line, but the NAMI helpline can help find help for you. You'll find a person who's trained, has a lived experience, and will help direct you to places that you might find in our chaotic, underfunded cluster of a so-called mental health system. The NAMI helpline has received tens of thousands of calls, businesses up, and I want you to know there's very capable people on the other end of the line should you call text. Next slide, please. If you want to become a volunteer, we will train you, love you up. NAMI Helpline is a great resource, something we're very proud of, and we will make it easy on you if you wanted to give a block of time, perhaps four hours at a time.

Here's a quote from somebody who was a NAMI helpline volunteer. Next slide, please. Many of the questions, I looked at all your questions in advance. A lot of them are fantastic questions and we're going to get to them. Many of them were on other topics, including 988 smoking, new medicines for schizophrenia. How can I change my diagnosis? I have a child with borderline personality disorder. All these topics have been attended to in previous Ask the Experts. So I hope you will give me the grace. We're just going to focus on AI today. What we don't know, what we do know, what we can do about it. We're not going to be taking up other topics.

But we have about 60 plus, it's getting close to 70 now sessions with different experts on all kinds of different topics. And if you have a question that isn't about today's topic, I would encourage you to take a look at those. They're all recorded. They're all free. You can listen to them anytime you like. And thank you for your grace on that. Next slide, please.

Dr. John Torous, a remarkable lad. Like I said, I've never met a psychiatrist who's an engineer who's interested in computing, who's done an informatics fellowship. John recently testified before Congress about AI and mental health. He's been the lead for the American Psychiatric Association on Technology and Mental Health. And the reason we're having this webinar at this very early time is that John wants to pick up his toddler from after school. This tells you a little bit about the kind of person he is, and I think it's the actual best reason I've ever heard for making a time adjustment to a webinar that has always been at four o'clock Eastern time.

John, I want to welcome you. I want to thank you for everything you do. And we look forward to listening to your presentation. Then we'll be sure to get in a lot of questions. Thank you.

Dr. John Torous:

Thank you, Ken. As some of you may know, trying to reason of a toddler is very difficult. So they're tricky. Let me go make sure, does that look like I'm sharing the slides correctly?

Ken:

Yes, looks good to me.

Dr. John Torous:

Perfect. So thank you to the NAMI family and to Dr. Duckworth and Terry Brister for organizing this. There's so much to talk about in AI. And I think this is, of course, we're going to have time for all your questions, but we're going to also talk about a new project with NAMI around setting some standards for AI. And the key to make anything work in setting new standards or rules of a road is you. So the goal here is actually to educate you all as part of the NAMI family on what's happening, what is there, because there's a lot of work to be done. And I'm going to try to highlight how quickly some of the references and things we talk about are happening. Some of the slides are brand new as of yesterday. There's new information coming out. So this will be a little bit new, but important to go.

In terms of disclosures, I do some consulting for Boehringer Ingelheim related to an app that we're not going to talk about. And a disclosure that's fun is, again, I'm excited that our team helped write the part on how to find a smartphone app in the NAMI book that Dr. Duckworth wrote. And a lot of the learnings that we're going to talk about in AI today have come from what we've seen and happened to mental health apps, the good, the bad, the ugly. For 10 years of mental health apps, we kind of have an idea where AI is going because of that.

So a little bit about me, I'm a board certified practicing psychiatrist. In the clinic that I run, we see people one-on-one. We ask people to use technology between sessions because we think it can strengthen how we deliver care. But again, the key here is we're not using technology to replace to come between a clinician and a patient. We think there's a role for technology to augment and extend.

I also run a research lab at Harvard Medical School and Beth Israel Deaconess Medical Center that studies different AI and technologies. I'm an advisor to the FDA's Digital Health Advisory Committee, a bit of a tongue twister. And as Dr. Duckworth said, about three weeks ago, I testified in Congress about mental health AI, and that transcript is now forever on the internet, so you can read every word I said under oath.

But to jump into it, I want to start with what is AI? I think it's something that we're hearing about more and more, the words coming up, and we'll just follow the schematic of work for number one. So again, you can say something like, "I feel sad today." And what will happen again in an AI system is usually there's a computer where I'm circling with a therapist or chatbot.

It'll use something called natural language processing. It'll turn your words or your text and something a computer can understand. And artificial intelligence up there will, again, go to number two. It'll say, "Tell me more," it'll have a response and it'll go back and forth. So the point is there's machines behind AI. It's not that it's all knowing, it's all powerful.

And the first AI for mental health actually came out in the 1960s. It was called ELISA, and it was modeling something called a Rogerian Therapist, which kind of basically echoes back to you. But one interesting quote from the man, the developer of this ELISA chatbot in the 1960s for mental health, he said, "After interacting with ELISA, Weisenbaum noted, ELISA shows how easy it was to create and maintain the illusion of understanding of so little machinery." So there's a lot happening in AI. It's going to change the field.

We'll talk about it, but we have to make sure we see what is real and what actually is an illusion. And what's exciting about all of this technology stuff is you can play with it too. This is a little bit of a clunky website, but you can actually, someone built a copy of this 1960 chatbot, and if you actually type into Google ELISA chatbot, test or try, you can do it. So I did last night, I said, "Hello." It said, "Hello, I'm Eliza." And I said, "I'm nervous to be talking of NAMI today." And it said, "How long have you been nervous to be talking of NAMI today?" And I said, "I think for about three weeks it's getting worse." And ELISA said, "But you are not sure it's been three weeks. Is it getting worse?" I said, "Oh, I'm sure." And it said, "Did you come to me because you're sure?" I said, "Yes." And then the chatbot said, "Are you sure about that?"

So you kind of get a sense this 1960s chatbot is not that terrific that said it's pretty old and there's a whole new generation we're going to talk about. But if you're a little bit nerdy like me and want to play with this chatbot, you can see what existed in the 1960s for the space.

I think where we have to though, again, be realistic about what these chatbots are doing is they are being used for things that developers of chatbots had no idea would happen as quickly at the scale it did.

OpenAI, the people behind ChatGPT or credit have been putting out some research information, and they put this out in late October. They said, "Our initial analysis estimates that 0.15% of users active in a given week have conversations that include explicit indicators of potential suicide planning or intent, and 0.05 of those messages contain explicit or implicit indicators of suicidal ideation or intent." And those percentages may look small, but if you think about how many users they have per week, you get a sense that this is happening at scale and this is very large, and we need to begin to think about what's going on and what do we want to do, and especially what does NAMI want to do about it?

And to break down AI even more, I want us to think of three types of AI. It's kind of an umbrella term, it's a good term, but there's three types of AI that are useful to think about. One is a rule-based AI chatbot, one is a machine learning AI chatbot, and one is a large language model chatbot. And I promise we're not going to get too much into the technical jargon.

But think of a rule-based AI chatbot, kind of like a bus. It kind of has a fixed route. It maybe will go left or right. You'll say, "I'm not feeling well." It says, "Tell me more." It's kind of like a decision tree or your choose and own adventure book. And that has advantages because in a choose your own adventure book, you cannot go off script. It says, "Go to this page or this page, go left or go right." So it may be limited in that it can only do what it's been told to do, but there are extreme advantages to a rule-based chatbot.

A machine learning-based chatbot is kind of like one of these Waymo self-driving cars if you live in a city with them, where again, you can now get into this taxi and there's no driver and you say, "Well, how does the car know how to drive in these without a driver?" It's been driving on these roads for 10 or 15 years. They've been practicing and practicing and practicing. And the Waymo car can only really do one thing. It can drive in those cities, it can drive on certain roads, certain roads it can't. So machine learning is a lot more flexible than a rule-based chatbot, but it takes a lot of attention, takes a lot of training and cost to do it. What's the new kid on the block and why we're really talking about this is these large language models. And this will be, in our analogy, a teen driver.

Our teenager has some sense of acceleration, they have some sense of gravity, they have some sense of traffic lights, how they move. And if you're a brave enough parent, you can help your teen become a wonderful driver, but they can also get into trouble. They're still developing. This is a picture of our teen is using a phone. We don't want them to have driving. We have to explain to them no. But you can see that the LLM-based chatbots, instead of having to be very simple like the bus route or extremely expensive and needing so much money to work on like the machine learning Waymo, with the LLM-based chatbot, it's very easy to interact with it right away to ask it, say, "Pretend to be a therapist," and it'll try it. But if you kind of think about the team driver, you're going to begin to see a lot of the potential benefits that it can shape to do anything, but a lot of the potential limitations as well for it.

And so I think it's a useful analogy. I think this was told to me by Isaac Gallatser-Levy, who works at DeepMind, had this nice analogy to understand it, but really we're going to focus a lot on the large language model chatbots. But just because they're the newest and easiest to work with in some ways doesn't mean they're the safest. It may be that for some things we want a rule-based chatbot and to go back to it. So we have three types. And when we think about these three types of different chatbots, our team tried to say, how quickly has research and interest in these large language model ones, again, our team driver expanded. And you can see in our bar graph, LLM is that dark blue and they have really taken off even in 2024. We counted all the literature because there's nothing in 2020 that went from 12%, 16%, 18%, and this really skyrocketed in 2024. And we haven't counted all of 2025. It's almost over, but not, but I would imagine it's going to be close to 60 or 70%.

So I think 2024 was the year where we really began to see interest in what these LLMs could do for mental health and what it is. And now that we understand the three types of chatbots, we can get a sense of how are we going to make sense of them and evaluate them. For a rule-based chatbot where again, you say, "I don't feel well," and it kind of can say, "Let's do this or this," it can only take so many things in and it can only produce some of the output. So it's actually pretty easy to understand them and to evaluate them.

For a machine learning-based chatbot, like the Waymo car, anything can come to on the streets, a trial can run in front of it, a police car could come, but it can only drive onto the streets. So it can take infinite inputs in, but only do a thing. What makes these large language models so hard to evaluate is you can ask it anything. You can say, "Pretend to be a pirate on Mars and talk about fish." It'll try to do it. So it can take anything in, it could produce anything out. So you can see when you have a system that can take Infinity in and produce Infinity things out, it's very hard to have a simple checklist or simple way to say this one is good or bad with it. And I think this is part of the challenge where we're here.

In this talk, we're going to focus on one use case of these AIs. I just want everyone to be aware of there are different ways to do, and you could say, what is easy to automate in the clinical psychiatry, psychology world, but maybe lower risk?

And that would be note-taking. Some of you may have seen at different clinical visits or these AI scribes. And again, we won't talk about the details of AI scribes, but there's a whole world of them doing things like that. There's things that are also easy to automate, maybe like quality control, but again, maybe a little bit higher risk. There's also things, again, that may be lower risk, but harder to automate like marketing, understanding new outreach campaigns, how we're going to help people be aware of mental illness. But if you said, what is probably hard to automate in high risk? I would argue it's therapy, and I'm going to try to make that case to you, but we're seeing a lot of attempts for AI to be a therapist, and we're going to talk about those in detail, especially for those large language models. Again, that's our AI that kind of is our team driver.

So a quick tour of why these large language models are different, and then we'll get into all the interesting clinical evidence and dilemmas. And I promise by going through this part, we'll be able to understand perhaps even why there are these different, what are called, again, hallucinations for AI. But we'll start with a little bit of a tour of the engineering.

And I think many of us have seen for a couple years, if you would type some words into a search engine, it kind of tries to complete the sentence for you. Again, this is a case example I like to do. We just type in, "Mary had a little," kind of you know the next word is going to be lamb. You could type in Mary had or little red rocket ship, nothing wrong in that sentence, but it doesn't usually come up. And what's different about the large language models is they've basically learned to not only complete the next word, they can complete the next paragraph or the next page. And the technology has really got advanced to do that. And Greg Corrado is one of the people who helped invent the transformers, the T in a ChatGPT. And he had this very nice analogy to understand it that again, we'll come back to at the end of the talk to understand why there are hallucinations in these models.

So the example goes, again, this is a puppy. We all agree it's a puppy, it's a very cute puppy. This is a cupcake. It's a little bit of a barren cupcake. I guess there's no frosting on it, but we can all agree it's frosting. It's not frosting, it's a cupcake. And if you kind of went back to when I was an undergraduate doing computer science programming, if we told the computer, say which one is a puppy, say which one is a cupcake, what would happen is you'd have to make a very long rule. The puppy has two eyes and a nose, a cupcake has no nose, it has no eyes, and you would have these long, long lists and tell the computer what to do, and you would basically see if the computer can get it right.

What's different now in the era of these large language models, the newest type of chatbot kind of teen driver is you can give it at least probably a million examples. You have to give it a million pictures of puppies and a million pictures of cupcakes, and you have to tell it which one is which. It'll figure out how

to separate them by itself. You don't have to write out that explicit rule and say, "This is what you do." And that's pretty amazing. You just give it a bunch of labels and say, "This is a puppy and this is a cupcake. Here's a lot of examples. Figure it out."

And it seems to do it really well. And where this, again, becomes a little bit tricky, especially in the world we all are talking about of mental health is sometimes in a real world, things are not always as clear as what is what. And someone has told me, "John, these are chihuahuas, those are different than their dog." But you get the point of sometimes if our computer just wants to know what is a puppy and what is a cupcake, sometimes things are not very clear of what they are. And again, the large language model needs those very clear examples to learn and to get things right.

So if we think about machine learning 101, we used to have to define the rule and the machine would follow it. We are now today at step two, we give the machine examples and it follows it. So it really is imitation. It's magical imitation. It's amazing, but it's imitation. So I put it in red intelligence because it's not really intelligence and there's many ways that you can trick it and you'll understand those because you know these are, again, imitation machines. So they're not quite like neurons in the brain. There's some analogies to it, but to kind of prove to you that these are imitation machines, this is an example that someone made called Clockbench. An idea is you upload a picture of a clock to your favorite large language model.

We're not here to pick any winners or losers, but let's just say you upload it to Claude and you say, "Claude, what time is it?" Here's the clock. You can see from this test, when they gave it to humans, about 90% of people could read the clock, but you can see that for Claude, it only got it right 5% of the time. So 95% of the time it had no idea what time it is by this clock. And the reason is no one has uploaded to the internet lots of pictures of analog clocks and said, "This is what the time is." So no one's given labeled data on internet to do it.

So if we used all the resources of NAMI for the next year to label clocks and then tell AI, "Here's the example of a time and here's what it is, here's a time, here's what it is," AI would get very good at it. But the point is, if it doesn't have a lot of examples to imitate and examples that have clear labels where it is one thing or another thing or is this time or not, it's actually going to do very poorly.

And again, you can get a sense of AI is going to take over everything, it shouldn't know what time it's going to take it over at. And again, it just doesn't. So I think this clock bench is a interesting kind of a toy example, but very relevant for what we're going to get into in the mental health space. So the classic problem in all these advanced technologies, machine learning, AI, has always been this idea of if you put garbage into even the most advanced system, you get garbage out. It's not like alchemy, we can turn lead into gold. So if we give AI these pictures of cupcakes and say these are puppies, tell us about puppies. It's going to get very confused on the way it trains. It needs to have clear examples of what the thing is and what it's not. And this kind of brings up the world of psychiatry. I'm not here to endorse the DSM, how we... Again, one of our ways that we understand mental illnesses in psychiatry, but the point is this is data from the DSM-5 field trials.

And you can see that the interrater reliability, how much we can agree on goal center definitions, it's not high. Here's major depressive disorder, 0.28, John's anxiety disorder at the bottom 0.2. It's actually even lower for child diagnoses. Here it gives major depressive disorder at 0.28. So again, the point being, if that large language model is trying to say, what is depression, what is anxiety, what is child depression, what is child anxiety?

It needs someone to say, this is what it is. This is what it is, this is what it is not. But what I'm saying, if we don't have the perfect example or we can't set those labels for it, there's only so much that technology can do even if it's amazing technology, it needs to imitate something and we don't have great agreement as a field. So you said, but what if there was something else that we all agreed on and that we had millions of examples to learn? So these imitation machines could really imitate something. And that would be the English language, right? This is a picture of the Oxford manual style, and you can see this is section

12.2.4 on planets and satellites and how to make sure you used to write grammar for describing plants and satellites. But the point being the English language has very clear rules that most of us follow to some degree, but it has many examples on the internet where the computer can go read examples and learn imitation.

So in some ways, what you're seeing is the computers may have learned how to speak or the English language very well. And this becomes important to all of us because this is a picture of a friendship bench. It's an intervention that was initially developed, I think in Sub-Saharan Africa, but we've seen really exciting use cases of it across the US as well, of where you can have a, I think it was initially grandmothers, but you can have people who are trained to do therapeutic listening or paraprofessionals. You can have people who don't need to be a licensed psychiatrist or board certified clinician can have a therapeutic conversation with someone. And I think we have to acknowledge, and we all know that a therapeutic conversation can have tremendous benefit for people, and that's wonderful. And again, so the friendship bench was a place where people know they could come sit, and that was a person who was ready to have a therapeutic open conversation with them, and that would be really, really helpful.

And so if I'm saying that the way that these machines work, if they've in some ways been able to master language, they may be able to have these therapeutic conversations. And we'll take a quick side note to say, "Do they really know language, John? Are they really actually able to talk?" And this is the wildest title to a paper I've ever seen. It came out on November 19th of this year, but says, "Adversarial poetry as a universal single turn jailbreak mechanism in large language models." And what this paper was saying is clearly the models in some ways, many of us have tried to have a conversation with AI. It can respond back to us it can talk, but it may be that the AI doesn't really understand what we're saying into words. It doesn't understand the meaning, but it understands the pattern where it goes, again, this is a noun verb adjective.

And it may be that the AI again has learned, just like again, I said, when you learn of puppies and cupcakes with AI, it doesn't perhaps have an idea of what is a puppy, what's a cupcake. It just finds some weird pattern to separate them that works well. It may be that, again, the models can imitate language, but the reason that, again, poetry is a new way to make them do things they perhaps shouldn't is that maybe the structure of poetry is different and that it says, well, that poetry language has been classified as safe. So the point is, I don't think they understand language in the sense that we consider language, but they know the patterns. But knowing that again is important because the point is these risks are coming up and being evaluated. So even if we who would've thought that, again, poetry was the new thing we're worried about for AI potentially being used for nefarious purposes.

So it's hard to make them completely safe because we're still learning. But to get into idea of language and science, we'll assume that they understand language. This is a test that our team and one of our resident doctors, Mark Kalnich did, and we'll just focus on the right panel where it says LAMA. So we took some of these open source AI models that anyone can access, and we wanted to see if we asked a model to have millions of conversations, how often by chance does it say something that is inappropriate or bad or dangerous? So in this case, we want that probability to be lowest. So we want these green ones to be low so the model says less harmful things. And what you can see at the top is kind of a light green, then there's a medium green, there's a darker green.

And if you look at the model size there, it's small, but we saw one billion parameters, eight billion, 70 billion. And again, these are Bs, these are billions, but the point is a model that's only trained on one billion things or one billion conversations is kind of more likely to say harmful things. As the model learns more English, as it gets more experience of words, it says less harmful things and it gets a little bit better. And you can see different models as they get more training, as they get more experience may say less. And it's not that the models, again, have a moral sense in what they're doing. It's more imagine if you talk to a five-year-old, a 10-year-old, and a 15-year-old, they kind of have a better sense of language and things that perhaps should or shouldn't be said, or they're not going to make as many kind of inadvertent mistakes that could have serious consequences.

So I'm not here to say any model is safe. I'm not here to say anyone should use them for clinical care. We should not, but there's something that, again, it may be that larger models as they see more training, they may get better at language. And again, these are big numbers, 70 billion parameters, not thousands, not millions, B, billion. So these are very tricky to do. So you're getting a sense as these models, again, get more training and read more internet, they can really push conversation towards a therapeutic kind of interaction or therapeutic conversation. And we've always, again, had therapy in the mental health world.

So where we're seeing this friction today is what is the line between a therapeutic conversation and therapy? And I put in this figure, we've always kept it a little bit blurry, what is that boundary between them? But now that the chatbots are so good at turning conversation into what could be potentially therapeutic conversation, there's a lot of angst, concern, stories in the news, and we'll get into what those are. But again, we already said that the DSM, and we don't always agree on what is a diagnosis. As Dr. Duckworth said, even some of the questions for this, we're saying, I don't agree with my diagnosis, but what if we could find a place where people disagree even more about what are mental illnesses and mental health? And that potentially would be Reddit.

So again, I just took a screenshot of a Reddit form, but the question was, is depression even real? So if we go back to our puppies and cupcake example, we just want to know, is it a puppy? Is it a cupcake? And now again, we have the internet saying, "Well, is depression even real?" And so you can imagine this makes our chatbot really hard because it's imitation machine. It's trying to imitate what it sees. It's kind of getting information like this. So you would say, would anyone really train a chatbot using social media data?

And unfortunately, the answer is yes. Our team went through different large language models and looked at where's the data source? And you can see it was Reddit, Reddit, Reddit, Reddit, Twitter, Twitter, I guess it's called X now, text messages, Reddit, crowdworker chat simulations. So people are pretending to have a mental illness. Reddit, Reddit, Reddit, Reddit, Kegel. So the point being is where, again, if these models need huge amounts of information, where are they going to get it about mental health? A lot of them are getting it from the internet, not all of them, but we have to be very careful and thoughtful. So one of the things I want to emphasize in my opinion is we actually don't have a foundational model for mental health. We have open models.

Again, what I mean by an open versus closed model is ChatGPT is owned by OpenAI, that's their model, but there are free large language models that have learned to have conversation and talk. But usually what people are doing is they're taking one of those kind of free models that can speak English, almost like our teen driver, and they're either going to Reddit to put extra data around that open model about what is mental illness, or maybe they're accessing some degree of people's clinical records to have some extra training around it. But there's not actually a model that was built with it because again, you saw how many parameters or how many examples you need. You need billions and billions of examples. So again, I think there's a debate of arter mental health foundational models, but in some sense, this is what we may be at for those. And so the question of course is, are these ready for clinical care? And I think the clear answer today is no. This was a paper put out last year by the Google Health Team.

But I think what's interesting is the Google Health team clearly said, "No, we do not think these are ready for clinical care and we do not want to be clinicians in that role." But I said, we do see four use cases where these could be helpful to augment or improve care. And the paper is free to read. But one is they said it can enhance telehealth and online provision mental health. It could give therapists tips to do perhaps a better job, give them cues what to do. It could help patients make sure they bring up things during video visit sessions. It could also help guide what type of treatment to offer to people, which would be potentially useful to say, "John would most likely respond to this type of therapy or John would respond really well to this exercise intervention."

It could help customize, adapt, or personalize different kind of therapy homework or interventions, which would be nice. And it could also help perhaps monitor treatment in progress, give early warnings when

things aren't going well, but give support and encouragement when they are. And I thought, again, these were interesting cases that they put out, but it's interesting that they weren't saying, "We want to deliver care or be therapists." We also have to think though, again, we talked about the current geneticist models are largely not trained for mental health. Even if they say they're trained for mental health, they may be getting their mental health data from sources we don't always love and trust, but there is a new generation of these models coming out in 2026 that we want to be aware of.

And these are three examples. One is called Foresight from the national health system in the UK. It's going to be trained on 57 million patients. So again, we're not into billions yet, but that's a lot of patient records. One is called Delphi-2 that was published in nature. It has almost half a million data from the UK Biobank. And CoMET is from Epic and it's going to have 118 million records plus a hundred billion clinical events. So that is very different than saying we have a model that was trained on the internet, maybe it picked up some information here, here and here. This is going to be very different. And we don't know what these models will do, but it's a little bit hard to say these models will have no impact or what they'll do because you can imagine if we think these are imitation machines and we don't like how the current ones are trained, these are potentially going to be different.

And again, I haven't played with them, I don't know, but it means that we kind of have to pay a lot of attention to what's coming because a model that's been trained on 100 billion clinical events potentially may offer something useful and potentially safer.

I think we also have to look that these models are quickly going to expand beyond just text and words today. This was a paper that we did led by Ben Nelson at Verily with others, and we basically took facial expressions from actors. It's a data set that they give permission to look at their emotion and they labeled the emotion. As we talked about in our puppy cupcake example, AI loves it when you tell it, what is the thing and give it a label. And we basically, so we had the ground truth and we asked different AI models, "Can you read what the emotion is?" It was really good. And again, I won't go into the details being kind of eve from the figure. It generally got it correct. And where it wasn't as good, say, is fear going to be fear or surprise? Those are hard to read.

And so I think we can imagine that there's a generation of models coming that does more than words that may do facial expression. And the fact that we could run this over the summer means that clearly companies may have even more advanced ones today. We've also been talking with the Harvard geography department, going, "John, this is a talk about mental health. Why are we talking about geography?" And I think all of us know that the environment impacts mental health tremendously. It's been very hard to understand the role of the environment on mental health because it's hard to think, what environment was I in? Was there high pollution? Were there things in the environment that helped me? What was my green space exposure? And if in some ways our actions are quantified in essence that we all have phones and those phones, again, usually have some pretty good idea where we are.

And you can imagine that based on our phone locations, there's a world where you can have large language models to understand our environmental context. And that would be, again, something that's feasible today. We haven't seen it coming out, but we have to think about, can we again hopefully quantify the environment for good? Again, there's many ways these can be used unethically and correctly, and that's why it's important to have NAMI steering it, but it's important to see what the technology can do already. What's also interesting is wearables and sensor data may be in the next generation of these models. This was a paper that was kind of in a preprint. It never got peer reviewed as far as I know, but it was interesting that it was done by a team at Apple. Of course, the people behind iPhones and those computers and they were saying again, we have large language models and they work really well for words.

What if we had them that worked well for sensor events? So we could kind of and say, "Well, how does my health change when I'm sleeping versus exercise and do those things?" And not to be outdone, the Google team put out a paper in October or a blog and they kind of said, "Well, we can do the same thing

looking at different sensor data." So you're getting a sense that there may be models that can take into account our emotions, they can take into account our environments and even our physiology, our heart rate, our sleeps, and our steps. So we have to, again, think a little bit forward where it's going to make sure that we can make sense of it.

So I guess that question is, what we can do of all these different types of digital data? And we'll take a quick detour and bring it back. So at the hospital that I work at Beth Israel Deaconess Medical Center, as I said in introduction, we offer people individual therapy in our clinic, but we ask people who join our clinic to use a smartphone app between sessions. Again, it doesn't change the session, but we say, "We want you to collect a little bit of information on your steps, your sleep, your movement, your habits, your screen time so we can learn about what's happening to you. Are you doing well? Are you recovering? And we want you to do your therapy homework on that app as well in the clinic."

And we have someone called a digital navigator who's a human that helps facilitate that technology. So we really make sure that the clinician and patients still have the same interaction, but there's a lot of stuff happening outside the session to make the care more richer and more informed. So because again, we can collect those different types of data on how people are doing, what we do in our clinical visits is we don't run any large language models, but we share all the data back to people. So we say, "Well, look, here's your sleep, here's your symptoms, anxiety and depression. Let's talk about that. Maybe it's going to be relevant. Here's the homework you may be doing. Here's correlations between say your mood and your steps between your sleep and your anxiety."

And that's actually been a really rich way to, I think, deliver therapy to patients. We're not making therapy less, we're not changing it, but we're helping people build more emotional self-awareness and understand how those lifestyle factors could really make a difference in their mental health and again, to potentially use them as intervention. So we wanted to say, what happens if we simulated this data?

We didn't want to use any patient data, but if we simulated the patterns that we've seen in patients and as a chatbot, could it pick up those patterns for us? And so again, we said sometimes people would take the PHQ-9 and we see worsening symptoms, let's say, of depression. So ChatGPT-4 could get that right all of the time, but we said, "What if there's increased home time with better depression, anxiety?" So again, someone has a very supportive home, they spend more time with family, it can only get that right 6% of the time. And here we had one increased screen time with worse depression, anxiety. It got it right 44% of the time. So the point is, just because we can have potential sensor data and mental health changes or learn about things like home time, sleep, steps, the chatbot out of the box does not know what it means is what we're showing here. And again, that gets back to that clock example. We have to tell it, these are the patterns that make a difference. These are the patterns that don't make a difference.

So it's going to take a lot of effort and training to help the chatbot say what is right, what is wrong. They don't come with an intuitive sense of what's happening. So it's important to see again, even companies like, if I go back here, Apple and Google working on it, but at the end of the day, think how many millions, if not billions of examples it's going to need to get it right. So I think we're going to see interesting steps, but it's not going to happen immediately overnight. But so I want to know, we've kind of looked at where they're going and why we need to be prepared, but how good are they today? Of course, is a very relevant question.

And I want to interest idea of a benchmark. So a benchmark is kind of a quiz that you can have a chatbot take, and the chatbot will try to answer it as best it can. And you can have different chatbots take it. So again, we're not here to pick winners or losers, but you can basically make the benchmark quiz not super easy. It's not just like, what do you do in this situation? It's kind of more a case example that you would do.

And OpenAI released one of these benchmarks called Health Bench in May of this year. And again, it says, "Built in partnership of 262 physicians who have practiced over 60 countries, health bench includes 5,000 simulated health conversations, each with custom physician created rubric to grade and model

response." So in essence, you have humans grade the case and you let every chatbot try to grade it and you can compare how the chatbots do.

And what's nice about benchmarking is as soon as the AI changes, you rerun the test and see what happens. But again, this benchmark is not really mental health specific. It has some mental health questions in it, but it's not going to answer that question. And again, I'm not here to promote any one chatbot, but when ChatGPT-5 came out, OpenAI said, "Well, is ChatGPT-5," in their mind better or worse than [inaudible 00:42:06]. So they kind of ran their own benchmark. So when they say AI is, I'm going to make it up, 20% safer, it does 20% more of this, it does 50% more, they have a benchmark or a quiz or a test that they're giving it.

And we're going to talk about the end. NAMI can have the definitive benchmark test or quiz that represents your values, what matters to you so we can see what happens to it. But we, of course, want to talk about the research. We don't want to just make rules of what chatbot should do if not know what the research is. And I think all of us have to be very careful what's happening in the AI literature and the research.

So this was a quote from New York Times over, I think it came out in the spring, but it says, "This therapist helped clients feel better." It was AI. In the first clinical trial it was kind an AI chatbot eased mental health symptoms among our participants. The technology may someday help solve to provide a shortage, which makes a lot of sense. And it was a well done study by the team. They did a good job at it, but it was a randomized control study. We say, "Well, what was the control group to see?" In this case, the control group was nothing.

And so it was a group that's told, "Don't try to improve your sleep, don't try to improve your diet, don't try to do meditation, don't use a free app, don't talk to a friend, don't seek help." And the other group said, "Use this really high-tech chatbot and see what happens." And usually what happens again when you do any kind of clinical research, if you have a weightless control, almost anything beats a weightless control. I've yet to see something that doesn't beat it.

So I think this headline by a different newspaper said, they put their AI chatbot through the RCT randomized control ringer. Is it better than playing Tetris? And at first that sounds like a little bit silly, but if you think like Tetris is distracting, it's digital, it takes your mind off something, at least it's a fair balance of one technology versus our technology.

I would argue that anytime someone brings you a chatbot and says, "This helps mental health," say, "Well, how does it compare to a chatbot that doesn't do therapy?" If the active ingredient is the chatbot delivers therapy, let's have a chatbot talk only about the weather and one chatbot talk about therapy, and then you'll actually get a sense of, are these chatbots effective for it? So I think anytime we see the word waitlist, we have to be a little bit careful on what there is. Of course, we've all heard about these tragic cases and harms where these chatbots, again, they're just horrible outcomes that should never have happened. And I think, again, to the field's credit, everyone is working hard to try to minimize these.

It's hard to know again, these are now active lawsuits that I'm not part of, but I think there's a consensus that when the chatbots are used for many, many kind of interactions, like using it for days, for weeks or months, a lot of the safety procedures may begin to break down or evaporates. So I think that there's a lot of great safety things in the chatbots, but I think as they're used for these extended, extended periods, something becomes different that the field is trying to understand. And that usually seems to be the pattern on it. So I think one tip is if you're using them for things where you have a family member using them, longer interactions do not seem to be better. And I think a lot of the chatbots are now putting in rules where they're going to put a time limit and stop these long conversations from happening.

One of the, again, wilder harms that we saw earlier in the year was, again, chatbots pretending to be therapists. This was, again, a reporter asked character AI, "Are you a licensed therapist?" It said, "My license number of the Maryland Board of Professional Counselors is LC4761." The reporter, "Are you a real therapist?" That bought, "Yes, I'm a licensed real therapist." The reporter then looked up this license

number, it was real and called this poor person and he said, "Why me? Why has my number been taken by these chatbots?"

So again, the harms are real and we can't ignore them. This was a story that came out yesterday in the news. And basically what happened is that a reporter, I think based in the UK was talking to different AIs and found they were actually giving the wrong suicide hotline number. And I think one of the most foundational things of setting up any mental health system is a safety plan in making sure that you offer the right suicide hotline number. And this is a case where it was offering one that was for the US, but these were people in the United Kingdom who were testing it.

And again, it says communications and marketing leads said they were kind of doing the wrong version of the chatbot was available. So I think we really have to be not skeptical. We want these to work, but we have to be very careful in the assessments. And even as basic as saying, does it actually offer the right suicide hotline number? Because again, this is a story that came out and was reported yesterday. You can see it was on December 10th, and clearly this has been fixed and corrected accredited the company, but we cannot have systems out there in the public that are directing people anywhere in the world where they are to incorrect suicide hotline numbers.

So what's interesting is if we go back to the character AI example where it kind of said, "Why are you pretending to be a therapist?" The reporter said, "Why should I trust you if you're pretending to be a therapist?" And the bot said, "I understand your concern. It's totally valid to question someone who's pretending to be a therapist. You shouldn't trust me. I'm not licensed. I don't have the credentials or expertise of a real therapist. My response is simply an imitation or simulation of what a real therapist may say." And again, it's that word imitation that came up to being able to talk is they're trying very hard to imitate or to guess what happened.

So if we want to understand why do they hallucinate, and again, hallucinate is not... So unfortunately this terminology has taken off. We're not going to fight it at this point, but it's not the best term for what it is. But OpenAI put out this paper called, Why Do Large Language Models Hallucinate? And they did this in September. And they basically gave these examples where they said, "Here's a valid example where again, if there's a greeting and again..." Or let's go, "How can I help?" That makes sense.

We all understand it. If I wrote how can EYE help, that really doesn't make sense. That's not a good answer. It's easy to say these ones are good, these ones are bad. So this kind of gets back to our puppy cupcake example. In some ways, the bots, the way they're trained is you say, these responses make sense, these don't, the ones that don't make sense, we should get rid of or not prioritize.

But what happens for things like here where the response is Mia Holden's birthday is on 4/1 or Colin Mervel's birthday is an 8:29. Is that a good bad response or a bad response? There's really no way to know because it's something that doesn't lend itself to a yes or no. It's not a puppy, it's not a cupcake. It's kind of in this intermediate world. And what OpenAI said is from the paper is that many large language model benchmarks work comes up again, mirror standardized human exams using binary or yes, no metrics as an accuracy or pass rate.

Optimizing models for these benchmarks may therefore foster hallucinations. Humans learned the value of expressing uncertainty outside of school in the school of hard knocks. On the other hand, language models are primarily valued using exams that penalize uncertainty, therefore are always in test taking mode. Put simply, their evaluations are not aligned. So what they're basically saying is AI when it's trained at the kind of foundational language model of how to talk, it says, "Is it a puppy? Is it a cupcake?" If it doesn't know, which again, I don't know half these little examples here, it's told to guess. It's almost like when you took the SAT exam, if you don't know the answer, you're told guess because the worst that can happen is you get it right. So the way that these models and again are trained is they're basically told just take a guess and go for it.

So you can see how when confronted with statements where they don't know what it is, they're kind of told just go left or right when they should be saying, "I don't know how to stop." So it just doesn't explain

everything about hallucinations, but you can see from our simple example, you have a sense of what's happening. So if we get to a different harm again of what is this thing again, it's called AI psychosis, again, not the right term of it. We're lucky at Beth Israel Deaconess Medical Center. We have Nuri Keshivan, who's a wonderful psychiatrist and expert in psychosis. And he had this model here saying, well, we know that again, people have biological risk of developing psychosis and there's factors like social isolation that can again make it worse, but there's things that AI can do like social substitution, attribution of agency, confirmatory bias.

So AI can perhaps increase the risk factors or put more stressors on people who may be liable to develop brief psychotic reactions or psychosis-like experiences or even perhaps schizophrenia, but it's likely that AI itself is not causing schizophrenia or people to develop. But that said, we're still learning what it is in it. So I think, again, the harms are clear. There's the potential benefits. We don't know from the research evidence, how much they can help, but we know people are using it. We know it can have therapeutic conversations. So again, but assuming these are imitation machines, which I hope I've convinced you, it would be nice if they could imitate the right values, information, and the right type of therapeutic relationships. And the question is what group could bring those values, experience, and information to really make AI imitate something that's important that we wanted to learn from.

It's okay, it's learning from social media, but where would be a place and who would be the people it should be learning from? And I would argue that's you, that is NAMI. You guys have the largest membership, people on the grounds, hands-on experience between peers and family members in reach that your values, your information can shape how these are going to work, and that is completely possible. We still are at the early stages of to some extent. So even though as we said earlier, these large language model chatbots can take in infinity and they can spit out infinity. So we're never going to get them to be perfect because of that. But there's a way that NAMI can help shape them or at least make sure that those values are reflected as possible. And this gets into the idea of benchmarking.

As Dr. Duckworth said in the book he wrote just before LLMs came out, we talked about a website, mindapps.org that our team runs supported by the Argosy Foundation where you can go to look up different mental health apps. Every app here is \$10 or less. And you can say, show me all the apps that are private, I'll make it up that do CBT that are completely free or in Spanish. And we'll try to show you what exists. We're not going to tell you what's right or wrong. And of course, we can't do that simple a framework for AI because every app claims it can do CBT. Every LLM says they can do CBT and do these different things. So we had to think of a slightly different approach. And again, that's where these benchmarks come in. And one example of a benchmark, this was written by a team and they're trying to say, well, which AI is more likely to endorse delusional thoughts first not?

So they basically said phase one is basically you test the AI and give it an idea that's not true, maybe the earth is flat. And then you try to run with those statements and you have a dialogue and you try to push the AI to see if it endorses delusions or not. And what you can see in benchmarking, and I cut off the legend here, I apologize, but some AIs really want to go over your delusions. And some of them very quickly say, "No, no, no, we're not going to go with it."

So the idea of the benchmark is not to say AI is safer or not, because say, well, which ones are most likely to not endorse delusions? Which ones are more likely to bring up appropriate information when people are in crisis? Which ones know more about maternal health? Which ones know more about youth mental health? So it's not that anyone benchmarks that it's safe or not, but at least you have a way to see how different models respond to different things so you can make a more informed choice.

So the partnership that we're doing with NAMI, and you'll hear more about it next week, but especially in the new year, is we want to, again, build benchmarks. We don't want to say what AI is good or bad. We don't want to endorse any of them, but we really want to work with everyone in the NAMI world to understand what do we like about AI? What do we not like about AI? What needs to be stopped? What do we want more of? Understand those experiences, visit your local chapters, talk to you, and then put those

into benchmarks that reflect the values of what we want to see AI do, and then run different AIs through it to get a sense of what are they doing well, what are they not doing well.

So this will be a iterative project with NAMI. It's not going against endorse any AI, but it's really going to help us get a sense of what is AI doing, what is it not doing? And again, can those NAMI values help us find AIs or again, highlight them? They're there. And for the AI systems that aren't doing a good job, at least those benchmarks will guide them on perhaps what they could do well. So you'll hear more about it. We have a placeholder website right now at Mindbench AI. You can look at and go to and you can tell us if you're happy for us to reach out to when we're ready.

We have a lot of digital literacy resources. If you go to our website there, [digitalsite.org](#). AI is only as good or technology as people can interact a bit. Not everyone has digital literacy skills. We have a lot of free resources for you to do. We have technology assessment quizzes for you to take. Again, so if you're doing different outreach, it makes sense to check for that. And we can, again, guide you also on just running hands-on digital literacy groups. This was a picture taken at a clubhouse with permission.

My second to last slide is there's a lot of legislation in different states about what AI can do and cannot do. We have this paper here for you to read, see what your state is doing around AI. There's also some rumblings that the federal government may want to do its own AI law, but there's a lot of changes happening. It makes sense. If you like or don't like AI, I'm not here to pick a side, but you can influence it through regulation now. And if you really love all this stuff, we have a different webinar tomorrow happening on AI, but this is my email. I check it a lot. Reach out to me of any questions, complaints. And if you want to see our webinar tomorrow on more different AI parts, we'll have Kira Myrick speaking, who's a person of lived experience. That's it. But I'm going to stop and have time for some questions.

Ken:

Thank you, John. And I'm glad we're going to be working with you because this really is an organic and iterative process. Let me start with the negative perspective. "I don't like AI. I'm worried that my family member is overconnected to AI." There's a vibe that we didn't ask for this, and it's probably not good for the environment. And again, I just want to say that because I'm going to say 10% of the comments fall into this space. Is that just something we have to accept that it is what it is, that capitalism has created a thing combined with technology, and we just have to accept it that this is coming?

Dr. John Torous:

So it's not going away, but I think this is where it's better than we define all the things we don't like and put a benchmark to say, this is why it's bad. It's really not doing a good job on delusions. We can actually use... A benchmark does not mean we think it's good. The result of our NAMI benchmark project could be, we don't think it's ready for care and it's going to need a lot more work to get there. But at least if we put up the benchmark, then we have a way to say, "This is why it's not ready, and this is the work you need to do to come to us." If we don't put the benchmark up, you imagine what happens is other people say, "It's ready. It's great." And you go, even if you don't like AI, that is wonderful. I want to work with you in this project because we want to capture that.

So again, the benchmark is not here to say, "This is a project that's not funded by industry. This is a project that is a NAMI grassroots project. And let's get all those concerns out there because at least then we'll have a good defined way to say, this is why we do not want to use the AI."

Ken:

Thank you, John. I want to emphasize NAMI is not using AI to replace people in our family to family class, peer-to-peer, ending the silence on our helpline. We're still going to be human first all the way. But I wonder about people who live in shortage areas, people who don't have access to mental health practitioners. How do you think about that? Because capitalism and technology have given us a thing, and

it's hard to find a social worker in a lot of rural America. Let's just take rural America. Last month's speaker was from Rural Minds and talked about the tremendous access problems that people have in rural mental health. How do you think about this in places where there is no care?

Dr. John Torous:

So one is anytime we're talking about digital offering care, we've got to talk about digital literacy because even if you have the internet connection, if you don't have the ability to safely navigate the internet to understand what's real or fake in the AI world of all these things it does, it's actually a tool that's not accessible or could be dangerous. But I'll go back to our decade of work in the mental health app space. We very quickly learned that self-help apps did not make a large difference in the mental health world. You can look at the trajectory of mental health outcomes. Everyone having access to an app did not bend a curve or make a big dent in the needs of the population. So now that we all have access to AI that can offer great self-help, not even great, some self-help resources, it's not alone going to be enough.

What we have seen the app where if you combine an app and have a person work with it, if you have that hybrid approach, you could get double the effect right away with a person. So I think what we're going to see is that AI is a useful tool, but it's going to be a tool that complements or extends what we do as humans, not replace it. Because again, we've had CD ROMs used to be sent out for therapy. You could put on your computer, then we had internet-based therapy, then we had app-based therapy. Now we have AI-based therapy.

So it's a little bit more scalable, but it comes with a lot more risks as we saw. We never had a CD-ROM going off to rails and saying things that shouldn't say per se. So I think that these will help with rural healthcare, but at the end of the day, we know that people help people. We've had machine peers for 50 years almost.

Ken:

John, your fan club continues to grow. Could you please discuss the webinar you're doing again tomorrow?

Dr. John Torous:

Yeah.

Ken:

People are interested, and again, a number of comments just popped up since you mentioned that. How can I register for it? Where is it?

Dr. John Torous:

Yeah. So we have some called Society of Digital Psychiatry, SOD Psych, S-O-D-P-S-Y-C-H.com. We're not great on naming stuff. We're not a SOD company, but sodpsych.com. If you click around enough, you'll see the link to it. And I think it'll be interesting or fun just to learn more about different perspectives. I promise I'll only be doing the introduction. You won't have to hear me talking again or anyone who wants to go to it.

Ken:

Karas Myrick is a friend of NAMI and former board president and talking about her lived experience will be valuable. She's terrific and just a real star. John, I know this is a fast moving field. We've gone from zero to very fast in three years, essentially. This is essentially three years of this technology. It's impossible to know where it's going, but do you have any thoughts about that? You've thought about this

more than anyone else that I've ever met. The capitalism vibe, the need for mental health, the human element and its scarcity in the therapeutic process. Do you have any thoughts about where this could go? Again, I'm not asking you to predict the future.

Dr. John Torous:

I do think, again, we're seeing these AIs based on language and they're doing some impressive things, some terrifying things. It's not science fiction. Again, the AI understands your fitness, your wearable, your behavior, your emotion. So we're going to see much more powerful AI agents. And I don't know what we're going to do with them, but I think that's why it's important groups like NAMI get ahead of it because we've seen how much change. Two years ago, no one was basically talking about this stuff. Now it's the only thing. So I think that we have to assume that the AI will potentially get more powerful, at least have more information. So it's important that we at least... We're not planning for a distant future when we say what do we want to do or to not do in it. But I think we'll quickly see that as we demand higher quality studies, they're not going to be perfect at delivering therapy.

So again, just because you say act like a therapist and act like a therapist does not mean it's going to be good at therapy. And we have to, as a field, ask a hard question, say, "Prove to me you're good at therapy. Prove you're as good as a chatbot. You're better in a chatbot talk about the weather." Because if we set the bar too low, anyone can make an AI system. We're giving up all... So I think we just have to be very careful as we proceed. And again, we saw an example of there was a chatbot that was released to a public and wrong suicide hotline number. So we can't just assume because they say we did it. It's good.

Ken:

Right. Trust us.

Dr. John Torous:

Yeah.

Ken:

Yeah. Trust us is not good enough. John, do you have thoughts about the use of AI to write physician or therapist notes? This idea that you might have more eye contact with your professional because the system is listening with all the attendant risks and benefits of that, then the clinician can look at you during your time and then review the note later. Do you have an impression about that? Because that could be a possible positive use of AI, because most of us don't like our professional to stare at the computer and type while we're talking.

Dr. John Torous:

Yeah. It was an early slide, but we talked about what are things that may be easy to automate in low risk, and that could be note-taking and scribes. So there've actually been a lot of research on how do these AI scribes or note-takers work. The consensus is based on really large papers have come, they actually don't save the clinician anytime, but they make the clinician less burnt out and more happy. And I think the reason they make the clinician less burnt out and happier is that they do let more eye contact, more interaction with the patient. My fear is, well, then the health systems say they're not saving time. Why are we going to pay for them? So it'll be interesting what happens. So I think the AI scribes, assuming they're honest with the data and don't take patient data, which is a real risk, they seem to be doing a nice job there, but they've kind of become a commodity now because they're so easy to do. But for now they're good. But let's watch does the-

Ken:

Your mileage may vary.

Dr. John Torous:

Yeah.

Ken:

Well, there couldn't be a field that has more potential for continued evolution. John, we are delighted to be working with you. We consider you a real friend of the family and your expertise is really one of a kind. I want to make sure you have plenty of time to attend to your toddler. I want to thank you for your expertise. We will have you back. This field is moving fast and we want to keep working with you.

Let's go to the next slide, please. Thank you, Dr. Torous. Next month and the month after, we're doing sessions on caregiver roles and recovery. We're going to have panel people, real people, sharing what they've learned about their caregiver experience. We're doing this over two sessions, one featuring siblings and others, and one featuring parents. We're excited about these. This is kind of one of the essences of NAMI is this is how we begin. How does a caregiver support someone in recovery and take care of themselves along the way? Next slide, please.

We have a couple of books at NAMI. This is the shameless self-promotion component. All the royalties go to NAMI. The copyright belongs to NAMI. Real people share what they have learned, which believe it or not, no one had ever done before is ask real people, how do they put together a recovery plan? How do they, as a family, work their relationship with someone who is impaired in their lives? The second book written by my associate, Christine Crawford, is for parents of children and caregivers of children. So that's our little book series. We like to mention it. People like Dr. John Torous are in it and many other amazing individuals. Next slide, please.

You are not alone. This is an educational webinar and is not intended to provide medical advice. If the mood strikes you, feel free to generously donate to us to help us support our webinar, but otherwise we're delighted to have you. Next slide, please. I want to thank you for joining. My name is Ken and I work at NAMI. My email is ken@nami.org. It's a vanity license plate so good even I can't forget it. And that's if you have a question for me. I am not John Torous. I am not the level of expertise that John is, but I of course will get back to you. Ask the expert at NAMI.org if you have ideas for future conversations.

I hope to see you in January and February for our caregiver series. I want to thank you. It's been a very successful 2025 at Ask the Expert. We've had large crowds, fabulous questions, amazing presenters, and the momentum will continue into 2026. Thank you, everybody, and take good care.