When Science Finds a Way

Season 4, Episode 2
AI and mental health: can we trust AI therapy?

Show notes

Episode description:

One in four of us uses AI chatbots or apps for mental health support, but could these tools have risks? In the second part of this special episode, Alisha Wainwright unpicks her own experiments with AI therapy, takes tips from Dr John Torous on staying safe when seeking help digitally, and hears about the risks of unregulated tools. But there's hope with these tools too. Professor Miranda Wolpert reveals how AI could help scale mental health interventions in the globally and even inspire whole new forms of therapy.

Transcript

Music starts 00:00

[00:00:00] **John Torous:** Again, it's hard to know the exact number of people say using mental health apps or AI chat bots. It's at least 25% of people. It's one in four. This is not rare.

Alisha Wainwright: Wow!

[00:00:14] **Alisha Wainwright:** Okay. Real talk. Have you ever used AI as a therapist? It's okay 'cause you are not alone. Many of us are already using these tools and I mean, I get it. It's cheaper. It's immediate. No long wait lists or awkward consultations, but as we heard last episode, it comes with some big problems too.

So we are back again with Dr. John Torous, director of the Digital Psychiatry Division at Beth Israel Deaconess Medical Center for our second episode on AI and mental health. And this time we are diving deeper into the world of AI therapy, chatbots with John and some fascinating contributors. We'll explore: How can AI help with mental healthcare's biggest challenges?

[00:01:08] **Miranda Wolpert:** The potential for AI to scale interventions to be able to reach millions of people across the world is enormous.

[00:01:15] **Grace Gatera:** I would be helpful to bridging the gap, especially for younger people who are more versatile, um, and live digitally.

[00:01:24] **Alisha Wainwright:** I was actually genuinely surprised that it felt like my most empathetic listening friend was talking to me. It was very disconcerting.

Alisha Wainwright: And what potential pitfalls need to be avoided?

Holly Coole: I don't think it's currently all that clear. For somebody that would like some support at four in the morning to know whether or not the product that they've happened to stumble upon on an app store is going to be helpful.

[00:01:58] **John Torous:** We've seen the rise and fall of mental health apps, we know where the story is going if we don't have clear rules of the road, so I'm gonna put it out there, it's anti-innovation not to want some type of oversight.

[00:02:10] **Alisha Wainwright:** Welcome back to When Science Finds a Way, a podcast about the science changing the world. I'm Alicia Wainwright, and you're listening to the second part of our conversation with John Toros on how AI is reshaping mental health.

Okay, let's get into it.

Alisha Wainwright: Hi John, and welcome back to the podcast.

[00:02:32] **John Torous:** Happy to be here again.

[00:02:33] **Alisha Wainwright:** Okay, so we touched very briefly on the topic of Therabots last episode. Um, so let's dive into that side of the conversation, but, uh, I hear you wanted to ask me something first, so I'll let you take the reins.

[00:02:49] **John Torous:** Thank you.

It's I, I'll be a host for only a minute, but again, there's so much interest in what are these chat bots for therapy. I was just wondering what is your experience or have you tried any of these chat bots offering therapy or therapeutic services?

[00:03:03] **Alisha Wainwright:** So in preparation for this episode and thinking about sort of my experience with my own mental health challenges.

And also my use of a therapist on and off over the years. I hopped on the, uh, AI chat bot service that I use for everything, and I just started engaging with it, having a therapeutic conversation. And I actually used two of them because I was curious how either would react - one of the Bots was actually really not good.

And it was kind of like, I don't know if you've ever asked, told a partner about your problems, and then they just try to find solutions. They're like, well, you should do this, this, this, this, and that, and you just don't feel heard. That chat, that chat bot was like that. And then another one kind of skirted the line between, I understand that you might be feeling these things, have you thought about this?

And I was actually genuinely surprised that it felt like my most empathetic listening friend was talking to me. It was very disconcerting because I wasn't prepared to think that it would do a good job. And it did.

[00:04:22] **John Torous:** Yeah. There's definitely something there to them. Right? That's why people are, are turning to them.

We, we could talk about some risks, but they definitely can offer some therapeutic elements out of the box today, which, which is frankly impressive. Right?

[00:04:36] **Alisha Wainwright:** It's impressive, but it's also kind of scary because

[00:04:39] **John Torous:** yeah,

[00:04:39] **Alisha Wainwright:** one chat bot, I walked away from that conversation feeling worse than if I hadn't spoken to it.

Off the jump and with another one, I walked away feeling genuinely better, and I felt a little bit more empowered to understand my circumstance in a new light. So I think it's kind of scary because I have the privilege of having access to two, uh, chat services that I pay a subscription for. But if I didn't and I just had free access to what was available, or most common, I might not necessarily be getting good therapeutic conversation.

[00:05:20] **John Torous:** Yeah, no, you, you can get some pretty harmful and horrific content.

[00:05:24] **Alisha Wainwright:** Well, I'm gonna take the reins back. Uh,

[00:05:26] **John Torous:** Sorry [laughs]

[00:05:27] **Alisha Wainwright:** It's my show, but I am gonna jump in with a question. So let's, let's zoom out very, uh, big picture. Just explain to me what exactly is a therapy bot.

[00:05:38] **John Torous:** Probably no one knows what a therapy bot is.

And I'll explain why these chat bots have become really good at language, especially the English language, because they have read everything on the internet and most things on the internet are written in English. And as these models have got bigger, we've gone from say, chat, GT three, to four, to five, they have become better and better at conversation.

And as we were saying, Alicia, we know that a therapeutic conversation can have some healing potential in, in some… it's useful… we, we all have a friend who can really elicit a great conversation, but when you have this therapeutic. Potential of conversation that can have healing powers for some people in some conditions.

It doesn't mean you're a therapist. It doesn't mean you know how to deal with emergencies. You know how to deal with subtleties. You know what the mental illness is doing. And what we're seeing right now in the world is what is that line between a therapeutic conversation and therapy? But those two have come uncomfortably close together to the part where it's almost impossible to figure out which is which, and that is… anytime there's a blurred boundary in healthcare, that's usually not a win for anyone.

[00:06:49] **Alisha Wainwright:** No, not at all. And I mean, how many mental health apps are there right now? Do you know?

[00:06:58] **John Torous:** We actually tried to count these in 2017.

[00:07:00] **Alisha Wainwright:** That's a long time ago, so I'm sure it's only gotten worse. Bigger, I, I don't wanna say worse, but bigger.

[00:07:05] **John Torous:** It was at least 10,000. And that's a number that we stuck with. We said we think there's 10,000 of them. And. It's a lot. And of course, yeah. Now what, eight years, almost nine years later it's going to be thousands and thousands of thousands of these, of these apps and it's just like, but the bots are proliferating so quickly too because there's a website, character ai, you can go on, you don't need any program experience, you clinical experience, you can kind of make your own mental health chat bot. It's not really delivering mental health, but you can call it Freud therapist bot, put it out there and offer it to people.

[00:07:45] **Alisha Wainwright:** Right. And how many of those apps do you think are using ai.

[00:07:49] **John Torous:** So, so we think about 20% of the apps now claim to have AI inside them. But the question is, what does that mean to have AI inside you?

[00:07:59] **Alisha Wainwright:** Yeah, it could, it could be used as, uh, like data analysis. It could be used as a chat bot. You're just, they're not very clear.

[00:08:11] **John Torous:** They're not very clear, and a lot of times it just means it has natural language processing. So just like you can call an airline and you can say, press one for ticket or yell one, and it'll do it.

So natural language processing uses AI to turn your words into ones and zeros. So I, I do worry or or wonder if a lot of the uses of AI in these apps are just kind of superficial.

[00:08:32] **Alisha Wainwright:** And do you think it's maybe like a bit of a buzz, you know, like, oh look, we're so technologically hip that our app uses ai, or is that an unfair characterization?

[00:08:44] **John Torous:** No, I think that's fair. 'cause if you're really adding AI in the sense that we think about today of generative AI in large language models. You're opening yourself up to a lot of risk, right? That these apps could start telling people things that are dangerous or unhelpful. So you have to have a very robust monitoring system if, if you're really offering AI in kind of the modern sense of the word.

So I think that the fact that these apps are not doing that tells us it's probably a little bit more superficial and trying to lead into the kind of buzz and have better marketing.

[00:09:17] **Alisha Wainwright:** I think it's interesting that John talked here about the noise, the marketing noise, the big promises, the failure of tools to make it clear what exactly they're offering, and it can be hard to get beyond that, especially when tools, at least at the moment, might be promising more than they can deliver.

But there is a reason why so many developers have rushed into this space and why AI carries that big buzzy punch that it does. And I think that is because there's real promise and opportunity too. So before I got much further with John, I wanted to share with him my conversation with Miranda Wolpert, director of Mental Health at Wellcome, who has a decidedly ambitious view of how transformative AI could be.

[00:10:12] **Miranda Wolpert:** So at the moment we have almost a billion people across the world who are affected by mental health problems that are holding them back from doing things that they want in life. And these problems generally start when people are quite young. They start in adolescence and they can derail whole lives.

They can stop people having the relationships they want, the jobs they want, they can even lead to premature death. And at the moment, for that billion people across the world, a tiny, tiny minority get the help they need. And that's for a range of reasons. Even where we have effective help, it's hard for people to find it, even when they find it can take up to a decade to find the help that's right for them and for many people, so some 90% of the population in global south, you can't access even something that may or may not help you. So there's a massive issue in terms of getting access to what you need, even for the richest person, but even more for the poorest person. And the potential for AI to scale interventions to be able to reach millions of people across the world is enormous. And also to address some of the access barriers we've got now, such as multiple languages, such as, um, cost, uh, and such as, uh, ability to get it at a time when you need it. So there are real things that AI can do that could transform access in the next way if it's done right, and if we take the safety concerns into account and if we think about, uh, how to do that in a safe and equitable way.

We really wanna make sure that what we don't do with this technology is bake all the faults and prejudices we have now. So what we wanna do is make sure this technology is developed from the start to overcome some of those issues rather than just exaggerating them. And we know there are real risks with using existing models as they are now. And there's real risks also of, of, um, not, uh, of intervening, not thoughtfully. So we know, for example, there's some examples of ai, which has been, we, uh, have actually induced psychosis in people or encouraged psychosis in people because it hasn't been used in a safe and careful way.

And we know that not all AI is, is built to be sort of therapeutic or helpful. On the other hand, we also know that humans, therapists are also not always helpful, and they also sometimes worsen conditions for people. So I think you need to take that in the round.

What I also think is really exciting is we're seeing completely different forms of therapy come forward, and I think we'll see more of those as we go forward that are not based on a model of two human beings sitting down and talking about the nature of the difficulties.

They may be based on completely different things. So for example, there's a set of researchers where, um. Funding who are looking at at work for people with psychosis who hear these very distressing voices and are learning ways to challenge those voices. And the way they're training people to challenge those voices is to have a sort of avatar of that person that they can sort of talk back to and learn how to speak back to at that moment… at the moment, that involves a human therapist talking through. But again, in the future, if that could be, uh, supported by ai, there may be ways of training people on how to address these thoughts, how to address, uh, these very distressing thought voices that allow 'em to get on with their lives. So I think they're gonna be a whole range of different interventions, um, that are going to be, uh, supported because we're going to be able to escape from a sort of

assumption that, of the way that what therapy means is gonna mean something very different within the next few years.

[00:13:33] **Alisha Wainwright:** Listening to Miranda talk about entirely new forms of therapy, including things like avatar based treatments, it really stuck with me. This is actually something we explored in more depth in season two of when science finds a way. So if you're curious, I implore you to go back and listen to that episode.

But overall, it's a reminder that therapy itself is already starting to look very different from what many of us might imagine, and it made me think about my own relationship to therapy and how much that's already shifted all without me really noticing.

Alisha Wainwright: When I first started going to therapy, this was uh, well before the pandemic.I would go in person and at the time I could never have conceived of doing a Zoom session. And then with the pandemic and the rise of having video calls, I don't even think twice anymore about engaging with my therapist on Zoom. So now I'm wondering, in five years, how open would I be to having an effective tool that functions like a therapist?

You know, I think my skepticism, I don't wanna say it's not reasonable. I think it's very reasonable, but I'm also really moved by Miranda's optimism because I do think there's a reason people are invested in trying to expand AI's use in the mental health space. I just, I feel like I need to hear more voices of the Mirandas of the world who are coming from it, from a medical stance and not from a, like a company stance.

[00:15:21] **John Torous:** Yeah, and I think what's exciting about Miranda's vision is it's very possible, right? Everything she said is reasonable and could work where it becomes different, and again, the transformation to telehealth and video visits, you're still connecting with kind of the same mental health provider or therapist.

You're just doing it now through a screen. What's different about this kind of AI revolution is you're gonna be connecting with someone different through a screen at a different entity, and that's where I think the risks become a little bit higher because you don't have that same, safety system, training level of what a therapist is.

Again, it'll be different and that could be potentially better, but I think that's where you want to really make sure you put the right guardrails and structure at the beginning because as Miranda said, right, we have a lot of biases and faults in the mental health system today. AI could certainly scale those up.

We, we've seen the safety issue, so it's almost, if we can launch it in the right direction today it could go in that path, but if we launch it in a, in the wrong direction, it could actually become very dangerous and, and really just a missed potential for the field. So, but I do think the entire mental health space has realized this is an amazing potential to act on, and it's kind of been the only word on people's mind for at least the last one, if not two years.

[00:16:42] **Alisha Wainwright:** You know what, I kind of feel a little similarly to autonomous vehicles. So I think when initially autonomous vehicles came to the forefront of, of conversation in media, people were like, no way.

You know, the idea that - what if they run over a person to save, like, you know, like how are they making ethical choices and et cetera, and now - I see them everywhere here in Los Angeles, um, and I've heard and read that they're fairly safe when you compare them to your average driver. And I think my initial skepticism of them have gone down as I've seen proof of their reliability.

And I'm just thinking to your point. Having that proof of reliability by setting regulations from now is how you can see, and I, I think it's much easier and visually obvious to regulate a car on the road. Right. Yeah. But it's a little bit more gray, would you say to think about how to incorporate regulations now, like for example, with mindfulness app or with a therapy bot app, what sort of regulations can we put in place to perhaps stop some of these apps that might actually be dangerous to users?

[00:18:04] **John Torous:** Yes, so I think the, the path towards innovation and division that you share, Miranda, is going to actually be regulation rules of the road, right? Because if, if we want to get technology to be like these self-driving cars. It took a lot of research and investment that was kind of done in the same direction. And if we have everyone pulling in a different direction, all saying their chatbot works today, it, it only harms the entire space and sets it backwards.

And the type of kind of regulation rules of the road we need are not strict. They'll be very transparent on how the bot works, what are the successes, what are the failures, and when they compare and say, I'm effective, they need to say compared to, what? They need to say am I, am I compared to a therapist, compared to nothing compared, to a bot about the weather. So as long as we have the facts and transparency, what I love about the scientific process in clinical research is things will get better quickly. This is not gonna take 15 years. Self-driving cars have been around for a long time. They're just getting better. AI is moving quick. It may not be ready to be an autonomous therapist in six months, but we could at least say, look, we all want to have some standards on what we do, what we report, we will either figure out works or doesn't work much quicker. But the risk is everyone says their chat bot is effective. No one talks about the risks and then consumers and clinicians are left trying to guess what to do. And what happens is then - patients, peers are smart, clinicians are smart - will say, we're not gonna use this stuff. It's too dangerous. It's too risky. So ironically, it actually hurts the entrepreneurs and investors.

[00:19:40] **Alisha Wainwright:** Yeah. You don't, you're leaving, what are they? They're throwing the baby out with the bath water. Yeah. You're not, you're, you're, you're just gonna say no to the whole enterprise when there could be some really good stuff in there.

Alisha Wainwright: John makes a really compelling case for smart, thoughtful regulation. So, how far away are we from having agreed rules of the road, and what are the people creating these safeguards learning about AI as they paved the way?

I spoke with Holly Coole, a former mental health nurse, now, the digital mental health manager at the Medicine and Healthcare Products Regulatory Agency, the MHRA. Holly is leading a welcome funded project that is looking to use regulation and evaluation to make digital mental health technology safer and more effective.

The project has a broad scope looking at digital tools of all kinds, but from her perspective, she does see a particular challenge when it comes to how to regulate those that rely on ai. Let's hear from her and get John's take.

Holly Coole: AI is, is something we've been cognizant of for a while, but I do think that the emergence of technologies has probably, maybe thought that people have suddenly, um, enabled us to, to kind of bring it into scope but it has always been a consideration. It has always had a regulatory framework around it is just the kind of awareness of that has perhaps been, um, low to middling up until recent years.

So there is a specific challenge with AI and that is inherently in, its, in its design in what it does, and that we can't always evidence how it's made the decision it's made or the outcome that it's come to, or the content that it's spat out in a particular scenario. So that presents inherent challenges in the way in which you evidence what a product does and how it does it. There are therabots or therapist chatbots that are constrained to models which can provide reliable, credible responses to prompts because they've been evidenced to do so. They have specific restrictive guardrails in place, and they are trained on clinical data sets.

There are others that range up to those that, as I mentioned, are maybe just. trained on data from the internet. No real clinical grounding in the data that's being used or the responses, and it's kind of just running wild.

Um, a user needs to know what product they are using in, in all honesty and putting it very simply I don't think it's currently all that clear for somebody that would like some support at four in the morning to know whether or not the product that they've happened to stumble upon on an app store is going to be helpful.

And when I say helpful, I mean helpful in a way that would replicate a clinical response to distress because. There's a reason why mental health professionals exist. There's a reason why the mental health system is in place, and it's to provide helpful clinical responses to distress and to mental health symptoms that could rapidly deteriorate.

And it's not just about language, it's about what's not said. It's about observation of that individual and recognizing and picking up on nuances of pauses in language or, you know, how how a person is presenting is a massive component of somebody's, um, mental state. And I just don't know whether or not users understand that something like that could actually cause more harm than good if it's not well evidenced and proven as safe and effective in those moments of acute distress. So I think more is needed to help the population and to help healthcare professionals understand that these technologies are available. I think actually everybody knows that they are available, but which ones are going to be most helpful in those moments?

So I think the the next priority for, for, not only for the project, so for the welcome funded project that that I'm working on, but also for the healthcare system, is to ensure that we prioritize education and understanding around these technologies to empower users to make an informed choice.

Alisha Wainwright: So I love that Holly's perspective came after Miranda's because they're pretty different.

And I'm wondering if you can speak on it. You know, Holly, who's tasked with regulating versus Miranda, who's just very optimistic about the direction AI could potentially take us. How do you, how do you reconcile their, their disconnect and, and can both things be true at the same time?

[00:25:25] **John Torous:** I, I think they both can be true 'cause what Holly said was, again, users need to know what they're using. So if, if you're making a medical claim verus a wellness claim, there's a very clear line in the sand. And there's nothing wrong with giving everyone access to a wellness tool. If everyone wants to try meditation or mindfulness and finds it's helpful, that is terrific and, and we can celebrate that.

But there's a big difference if you now claim your wellness tool is a medical tool, and as you heard, there's no clear, it's unclear what you're actually getting, so it's almost where the difference is, as Holly's saying, if you think your chat bot and AI is so powerful that cures a medical illness, prove it.

Show us that she's not saying don't do it. She's saying. Put your, put your evidence where your claim is and Miranda saying, we're gonna help build that evidence, we want it to get there, but Miranda is saying we're open to it. But even Miranda said, look, we see tremendous risk of bias, of faults in the mental health system being scaled up, of safety issues, we've all seen those lawsuits, so I think that they're actually just saying two parts of the same thing - we, we want to get AI to do as much wellness as possible, and we want to celebrate the ones that can cross over into the medical realm, but we don't want to let everything make a medical claim because then as Holly said, you don't even know what the heck you're using.

At some point it just becomes a, a forest, a a of mess and, and the real risk is then we lose trust as we talked about. And as soon as you lose the trust of clinicians and patients, you've kind of destroyed the potential of this new technology. And as we know, we don't get that many new technologies and things in mental health.

Like it's an amazing opportunity, but we, we need to get it right.

[00:27:09] **Alisha Wainwright:** I also think about, again, as a, uh, from a patient's perspective, a consumer, a user. I did not know until researching this episode that there was a difference in terms of a medical device versus a wellness or mindfulness app. I honestly thought they were all in the wellness space, and so as a consumer, I also feel frustrated that I don't even know.

What my options are, and you know, for many people in the world, they might have mild symptoms of anxiety or depression. You might engage with a therapy bot, but if you have some familiarity with a therapy bot, but you don't have a therapist, and then you experience an extreme mental health crisis or something that needs that crossover, do you know not to use something you've been accustomed to using?

Do you know what I mean? I, I, I just wonder, that people don't necessarily know how to make that distinction when they're going through a, a crisis themselves.

[00:28:13] **John Torous:** It, it really should be that these wellness things call themselves wellness and upfront about what they are. And, and what happens is there's kind of this, we said there's this blurred line always between is - it stress or anxiety? Is it mood or depression? And, and again, what's unfortunate is there's a lot of, again, yeah, there's always gonna be people to push the boundaries, right? They're gonna kind of try to make their technology look more appealing, try to, to kind of push it. And, and we do have to, in some ways understand what is and what can, so people know what it is, but it's very hard as, even a clinician who studies this to look at these websites and go, is that one actually

offering therapy or wellness? I, I cannot tell. And sometimes, the only way I can tell, and I hate to say this, I have to read the terms and conditions because legally it says we offer no psychiatric help, no medical help, if you dare have a crisis, you violated terms and conditions, you need to get off our service immediately. And you go, whoa, that's very strong language. So sometimes if you're considering, again, I don't recommend these for treatment at all, but if you're considering using one, talking to one, look at what they claim in the terms of service, because that's where you're gonna figure out what the company really says. And every single AI chat bot that I've seen in the world, we will disavow any clinical claims or medical liability into terms of service. So the question is, if they're not willing to kind of legally put their, their put their kind of words behind kind of risk - that that's kind of all you need to know and it's hopefully will get better but that's where it's nice to have people like Holly who are saying, look, I want to find and celebrate the ones that can cross that boundary, but I also want to stop the ones that are not ready to be there too, so that we don't get confused.

[00:30:01] **Alisha Wainwright:** Sure.

Alisha Wainwright: So there are people like Holly and John, clinicians, policy advisors, regulators who are working to create shared industry standards that hold developers to account.

But there is another key perspective that needs to be considered. The users themselves, those perspectives are key to creating realistic, practical guidance that meets the needs of those who need those tools the most. I spoke with Grace Gatera, a lived experience advisor who is working with Holly, about her role in the project and why, including lived experience perspectives is going to be key to getting this right.

[00:30:49] **Grace Gatera:** My per my family and I are survivors of the 1994 genocide against a Tutsi, um, and living as refugees. Uh, in Uganda and elsewhere, we experienced a lot of mental health, uh, challenges that arose from our experiences with the war, but also just the hardness of living as refugees in a country that's not your own.

Um, I remember being diagnosed with all sorts of things, however, care was not present and it was not consistent, especially at a time before like digital help was available. And so when I started receiving care, which was very recently, as recent as 2018, a combination of things helped me. But, uh, a big thing that helped was finding community, uh, digitally and also finding support digitally because, um, as you may know, as Rawanda is, is really trying hard and has helped the survive, the survivors get support, get mental health care. But, uh, we're still under-resourced and there's still very few psychiatrists and therapists for a population of 13 million. These digital tools and these digital mental healthcare, uh, would be helpful to bridging the gap, especially for younger people who are more versatile, um, and live digitally more than our population, our generation. Uh, a lot of young people are going to AI for questions, for support to heal their loneliness, and they're using it as a tool in their everyday life in a way that has, I don't think I've ever seen before, and so with such power, of course, comes great responsibility. We don't have an idea of how unregulated the data sharing is, especially for ai, for the AI space, um, and we're still learning about AI itself. Uh, governments are just now starting to think about AI regulations for populations, especially since some cases are coming up about, uh, misuse of AI for mental health care leading to drastic situations. So with welcome as a consultant, I was brought onto the MHRA and nice's collaboration with welcome to think about regulating mental health. Tools, um, and digital devices. Um, and we truly believe that lived experience should be part of it because we are the service users.

Uh, we are the end result of this work, and we are the people who have been using these digital tools and devices for a long time. So having, I think having lived experience, expertise embedded into from the very beginning, I think will provide direction and will spare sometimes expensive mistakes or misdirections just guiding us directly to what the core problems are and how to solve them.

I do think it can be hard, I want talking about in my past, but I think I share with a lot of lived experience advocates and advisors, uh, in this space, which I also think I share with the rest of the team and all the people who have poured their efforts into regulating digital spaces, uh, is it may seem negative, it may seem that we're trying to curtail things or, or like, you know, be wet blankets of sorts, but I think it's, it's passion for the future. We know the power that goes into digital tools. We know the hard work that people pour into making sure that mental health care is accessible, is accessible. And also we know the power of AI as it is. It's sweeping the world. It's changing the way we work, it's changing the way people access, you know, the internet and work on it. So, um, I think this is just the first step in sort of regulating digital tools, especially for mental health. But I think it's just creating a little bit of a beginning for what I hope is going to be a widespread movement, uh, taken on by, by governments and policy makers across the world.

[00:34:57] **Alisha Wainwright:** So I'm curious, why do you think it is so important to have people with lived experience involved in the regulation process?

[00:35:07] **John Torous:** Especially when we talk about kind of the AI in these kind of therapy roles, you say, where did it learn all the information about therapy? It kind of went to Reddit and social media and scrape it. So it took all the information from people of lived experience. Who is it trying to serve directly? It's not trying to give therapy to therapist, it may in part, is trying to then actually directly serve people. So in essence, it's taking all information from people. It's geared towards people. So, so it's only logical that the people it's designed for are, are part of the design. And the people that you could argue it took or stole information from are part of design too because imagine a world where it doesn't go to just Reddit and read forms about mental illness. It talks to people of experience that learns from what matters to them and then it learns from them how to best deliver that care back. Like we, we could actually make what we see today more personalized, more effective, and based off better information.

[00:36:05] **Alisha Wainwright:** And I love how self-aware, uh, grace was to say that she's understands that she might be perceived as a wet blanket, but I don't know. What would you say to someone who says that regulation might stifle innovation in this space?

[00:36:22] **John Torous:** I, I think it's actually anti-innovation to say you don't want regulation, right? To say you just want to be chaotic. I, I appreciate where she's coming from of that narrative, but really again, we've seen the rise and fall of mental health outs. We know where the story is going. If we don't have clear rules of the road and. So I'm gonna put out there, it's anti-innovation not to want some type of oversight.

[00:36:44] **Alisha Wainwright:** I, I agree. I think, you know, the technology space is sort of founded on this concept of move fast and break things, but we're specifically talking about people's mental health, people's lives, and I think you can't move through the space in that same way and expect there not to be really negative outcomes if there aren't any regulations or oversight.

[00:37:09] **John Torous:** You can still move fast and be rigorous too, right? It's, it's, there's almost this implication. You don't need to be slow. You can do massive decentralized studies. You can test this stuff ethically, and you, you may not, you may be 10% slower, but you, you're not going to kind of have a, feels like this 15 year bench to bedside to develop a new medication.

No one thinks it's gonna take 15 years to kind of develop a new. Chatbot. A a a young person can go to character AI and make their own in a couple minutes,

[00:37:37] **Alisha Wainwright:** Which is wild! So in thinking about the 10,000 mindfulness, uh, apps out there, if I am in the market for a wellness tool, a medical device tool in the form of an app, do you have any tips on what I should look out for and consider while I'm browsing the app store?

[00:37:59] **John Torous:** Yeah, so one is there's a lot of these things, so you can be a picky consumer. You don't have to pick the one that you see first that has the most marketing and be skeptical, and you can almost test drive them out, right? A lot of them are free to download or have free trials. There's no reason not to give them a shot.

The evidence actually supports, sometimes the ones that work the best are the ones that you stick with or are most engaging for you, which is different than your friends, your family. Even you and the past are future. So give them a test drive and we have a website, mind apps.org. It doesn't index every mental health app, but we have read the privacy policy for you, so you can kind of say, show me the ones that have a privacy policy. Show me ones that don't. And if you're in doubt, clinicians would love it if you said, here's an app I'm considering, I'm using - don't be shy about bringing it up into clinical visit. Clinicians know these exist. They know AI exists. There's no harm. No one's gonna judge you by saying, Hey, I tried this chat bot. What did you think of it? I tried this app. What do you think of it? I, I think it's important that we all just tell each other what we're doing, and I promise you clinicians will be excited, if your clinician's not excited, contact me. I'll contact your clinician. I'll be gonna promise.

[00:39:11] **Alisha Wainwright:** I, I, I'm really glad that you highlighted this sort of trepidation you might feel on talking about it because. In preparation for this episode over dinner a couple nights ago, I mentioned that I was using an app to engage in therapeutic conversation, and my friend was like, well, you know, so and so does that too. But the way she said it, it was kind of like, you know, it's a, a sort of a hush hush thing sort of in the way some people might talk about going to therapy in general, you know, it's a little stigmatized still, even in a, I would think, a pretty forward thinking community that I, that I moved through it. It is still a little, I, I'm curious if you have any thoughts on that,

[00:39:52] **John Torous:** we've, again, it's hard to know the exact number of people say using mental health apps or AI chat bots, it's at least 25% of people. It's one in four. This is not rare. Wow.

[00:40:02] **Alisha Wainwright:** Wow.

[00:40:02] **John Torous:** Okay. It's really common. Mm-hmm. It may be up to 50% have at least tried at once to done it. So what I at least tell clinicians is again.

If you're not asking people what they're using of these tools, you're kind of missing half the… like half the people are gonna have it. So I, I think the stigma still exists, but it's a lot more common than all of us think.

[00:40:22] **Alisha Wainwright:** So looking across everything that we've talked about, yeah. In both of these episodes, what is the biggest opportunity in mental health care with ai and what is perhaps the biggest risk?

[00:40:35] **John Torous:** I, I think the biggest opportunity is really gonna first start in using these tools to improve psychoeducation and resources and self-help around mental health. You could always go to the bookstore since probably the printing press exists and get a self-help book and go through it. That was effective.

Regulators were happy of that. People are happy… but it was hard to get through self-help books. They don't work for everyone. But there is a world where, without, again, getting onto clinical side, we can use ai, to really improve access to good quality information, personalized information, skills-based things that don't cross into treating a mental illness, and there's no reason not to deliver on that potential if we can do it safely.

I think the greatest risk is again, losing the trust of patients and clinicians by over-hyping the space, by making it so dangerous that regulators have to go far beyond what Holly's talking about. She is just saying, I wanna know what's good and label it, but if it becomes so dangerous that it becomes a political issue, that people are afraid of it, that's where you're going to have a different issue.

So I think that we have to be. Mindful in what we say we… it can do. And if we're gonna make medical claims back those up because if we do push it that way, we will break things, we'll harm people and we will miss the opportunity to have this work as we all want it.

[00:42:00] **Alisha Wainwright:** Well said. Thank you so much.

[00:42:02] **John Torous:** Thank you so much.

[00:42:04] **Alisha Wainwright:** We have come to the end of our two part deep dive into the world of AI and mental health. Thank you to our intrepid guide, Dr. John Torous, and to our contributors, Miranda Wolpert, Holly Coole and Grace Gatera. They all shared such thoughtful perspectives, and I've come out the other side with a much more nuanced picture of this very new frontier.

I can see how it's not just about techno optimists versus technos. Skeptics. It's about how we can make sure not to waste the opportunities AI offers by moving too fast without proper research or regulation. When tools claim to offer more than they can deliver, we all lose out. If you enjoyed this episode, make sure to follow us wherever you get your podcasts and like, and subscribe on YouTube when Science finds a way is brought to you by welcome, a charitable foundation that supports the global scientific community working to build a healthier future For everyone.

To learn more about the work they do, visit their website welcome.org/podcast. That's welcome with two Ls.