

Safety Tech Podcast [Episode 4] – Breaking the cycle of hate and harassment: talking to the people tackling online abuse

Tue, 4/19 11:26AM • 36:09

SUMMARY KEYWORDS

online, metaverse, technology, content, behaviour, people, models, problem, harm, violence, arwen, safety, moonshot, organisation, human, engaging, platform, data, world, companies

SPEAKERS

Matthew McGrory, Nina Jane Patel, Ben Whitelaw, Vidhya Ramalingam

Ben Whitelaw 00:01

Welcome to the safety tech podcast brought to you by the safety tech innovation network. My name is Ben Whitelaw, and I'm the founder and editor of everything in moderation, a weekly newsletter dedicated to online safety and content moderation. Today, we're focusing on violence, hate speech and harassment, and showing how online safety technology can be used to fight back against it. Some of the experiences described in this episode may be triggering for some listeners, thanks for joining us.

Nina Jane Patel 00:50

So over the past 18 months, I have been exploring a wide range of digital spaces in VR observing and conducting research for my work.

Ben Whitelaw 01:02

Nina Jane Patel is a doctoral researcher of the metaverse investigating the physiological and psychological impact of spending time in immersive virtual environments.

Nina Jane Patel 01:12

This past December, I entered a new experience available on the Oculus quest 2

Ben Whitelaw 01:19

The Oculus quest for anyone that doesn't know is a virtual reality headset developed by Oculus, a division of Facebook, which allows the wearer to interact with digital environments via an avatar

Nina Jane Patel 01:30

to access it you have to log in with your Facebook account, and I chose an avatar that I feel resembles me as I look in the organic physical world. Within 60 seconds of entering the meta horizon venues

lobby area, which is a kind of shared communal space, a social space that resembles the lobby of a movie theatre or a conference venue. Three to four male avatars with male voices surrounded my avatar, me, and came very close to me started harassing me, verbally, they were harassing me and also sexually, making sexual advances saying sexually inappropriate words and they touched and groped my avatar. While they took selfie photos within the virtual space. I asked them to stop repeatedly. They didn't. I tried to move away within the lobby area, but they followed me and continued to verbal harass me and extend sexual advances towards me. They were laughing, they were very aggressive. And they were quite relentless in their, what felt like very much an attack. Then I went into a flight or fight and I actually froze. And it felt quite surreal. And the only way I could effectively end the experience was to take my headset off. Even as I took the visual headset off, I could still hear them yelling in verbally harassing me. Part of me just wanted to forget that it had happened. But as the more I thought about it, I realised that the experience could have happened to anyone. And in you know, for relating that to me that could have been one of my children who just put on the headset and think that this is a fun interaction. And then in fact, it's it's actually quite a damaging interaction.

Ben Whitelaw 03:32

But so decided to find out if her experience exploring this virtual world was typical.

Nina Jane Patel 03:39

And so, uh, you know, I originally posted about this, you know, an online social platform, the community of Metaverse researchers. And what was alarming was the amount of hate that was posted in that group in response to my question, my question was like, has anyone else experienced this? And that kind of a plethora of negative responses about how stupid I am. Don't choose a female avatar, it's a simple fix, or I get a lot of words about how stupid I am to think that this was real or that I should even justified feeling like this was an attack. And then lots of kinds of jokes about oh, well, I've been murdered in VR millions of times or you know, all these kinds of relating to games that people have experienced violence, but they choose to that's an environment that they choose to go and shoot zombies or steal cars, or whatever it is, but I entered this space as a social environment, and I certainly did not choose to be verbally and sexually harassed in this space.

Ben Whitelaw 04:43

Virtual reality might represent exciting cutting edge technology, but what Patel had experienced was all too familiar to women and marginalised groups. According to research published last year, almost one in five women have reported being sexually harassed online and LGBTQ plus people, women of colour, and black women in particular, are often disproportionately targeted with other forms of online abuse. In many cases, this abuse violates their human rights.

Nina Jane Patel 05:11

The Metaverse is a new worlds and they're old problems. So what we're seeing is a continuation of of issues that exist in real life and online and all of our digital spaces. And what we know about the evolution of the Internet, you know, since 1995, what we prioritised from day one was anonymity, over accountability in terms of our actions on the internet. And then governments key stakeholders were very slow to take up any sort of action, because there's always the question of freedom of speech. And the question is, can we learn from these mistakes of the Internet where we know that damaging

behaviour is happening. And I think that we can learn from our mistakes with the Internet. And we all have a moral and ethical responsibility as designers and builders of this future world of the metaverse to take action now that will prevent an abundance of mental health issues and potential damage and trauma to the next generation of children who are today integrating this technology into their lives and on a daily basis, interacting socially on these platforms.

Matthew McGrory 06:30

I did a bunch of research and I sort of couldn't really get my head around why there weren't more companies in what I now know to be the trust and safety sector.

Ben Whitelaw 06:41

One person who has become aware of the size of the internet's hate speech problem is Matt McGrory. His company Arwen is one of the hosts of safety tech providers trying to combat toxicity by giving users greater control over what they see online.

Matthew McGrory 06:55

And it was very surprising that a problem that was so prevalent and so talked about in the media didn't have a whole host of companies trying to tackle this properly, this was about 18 months ago. And really, that's where this kind of embryo of kind of Arwen was born. And also, from my, I suppose, pandemic frustration, I have three children, 15 and 12, and 7, and trying to implement safety tools for them to use during homeschooling was just so challenging. I went through about four tools for my son, every time I put one in, he got around it. And it's not about mom and dad trying to block their behaviour or spy on them. In the fullness of time, we want to have an Arwen that sits on the phone. And just lets us know when bad things are happening. It's not going to spy on your conversation. But we kind of know that some bullying is going on, or there might be some self harms content being read out, we highlight those problems surface them so they can be dealt with in the real world.

Ben Whitelaw 07:57

Can you tell us a bit about Arwen and what it does?

Matthew McGrory 08:02

So Arwen is a tool designed to remove unwanted content from your public channels, specifically social media, like Twitter, Instagram, Facebook, and we've designed it to either work in the background for individuals or for businesses and brands work as part of a human based moderation team.

Ben Whitelaw 08:25

Tell me a little bit more detail if you can about how the product actually works. Technically speaking,

Matthew McGrory 08:30

Yeah. So I tend to cut this down into four parts. So I call it collect, classify, manage, and act, the kind of four things we try and do. So we collect your data, which could be in this case, we're talking about social media, we classify it against AI models. And then we do something in the moment, which we call manage, which is kind of we deal with that content in the way that the user wants that dealt with in that particular moment. And then we have Act, which is when certain thresholds are met, like someone

might cross over the threshold into illegality, you might want something to happen in the future, you might want to take legal action against someone, you might want to write them a letter to say, Stop doing this, please. There are different ways that different organisations and individuals react to what's been put in front of them in terms of the content.

Ben Whitelaw 09:21

And it acts in some respects, like a filter, is that right? You can understand what they're receiving, the messages they're receiving, and then screen them accordingly. Is that right?

Matthew McGrory 09:32

That's absolutely right. So there are various forms of word filters that are already in place. There are tools used by Twitter has at all, you can put in your own words, and it will filter out certain words. The challenge with that with certain topics like racism, a good example is the word black. You could be talking about someone's black tie, or you could be using that word to cast insult and someone in a racial slur. What we do is we denote the differences between those two scenarios using the AI models, and we put that into a traffic light system, so green is safe, red is severe. And this is stuff that we really want to get rid of. And then we have an amber, which is a suspect queue. And really what we're trying to do is get the green and red queues to be as intelligent as possible. So the human element, the human moderation, where the human has to step in, is as small a queue as possible. And is for those kind of nuanced examples. What we're trying to do is make sure that we put the right stuff in the red queue in the right stuff in the green queue as often as possible, because that's really useful to people that are managing this

Ben Whitelaw 10:41

is the idea that platforms don't need or services don't need humans at all, to moderate content. what's your what's your take on that?

Matthew McGrory 10:50

I'm very much against humans not being involved. You know, I was brought up on Terminator movies where the machine takes over and rules the world, I don't think we're quite going in that direction. But it's a similar concept that human beings need to be masters of the technology in my view, and we need to have a control over it, we need to understand what we're trying to achieve each time, we implement some technology. And we need to assess the consequences, you're always going to need human moderator for nuance. Humans, ingenious in positive ways, and also in horrible ways. And they will continue to be so. I've seen Michael Johnson, the 400 metre runner, as a specific bit of abuse, they abused him racially in text. And they did it using emojis. And then they did a picture of him winning the 400 metres at the Atlanta games, and then turning into a monkey, it was just like the time taken for someone to kind of go through and do that. And I just find it staggering that someone spent the time to do that really is flabbergasting. But you're going to need a human moderator to kind of - I mean that's fairly obvious - But people are doing more nuanced things. So we need humans to be there to kind of be the guardians of ensuring what's right, what's wrong. Which one do we put in each bucket?

Ben Whitelaw 12:11

The use of artificial intelligence to mimic human decision making raises some big ethical questions when it comes to online safety. And, Matthew, in terms of the data itself, how are you ensuring that the data that you're using to classify the materials that you find on these platforms? How do you ensure that the models themselves are okay? Because often, we've seen in a number of cases this past year, models are built by humans. And they often leave out other groups and they reflect our own biases. How do you deal with that?

Matthew McGrory 12:47

So this is sort of one of our uniquenesses. So we actually utilise numerous AI models, and we blend them together. And there's obviously bias within the blending. But that puts us in a quite a unique position in that. So we will work with a number of supplier partners. So the minute for our kind of anti hate platform, we've got six different AI models feeding into that. Because we've recognised inherently, an individual model is biased, you know, there will not be one model to rule them all. It's just impossible. And speaking to all the professors and data scientists I've spoken to what quickly became evident is that there are literally hundreds upon 1000s of different academics working around the world, on different ways of viewing this problem. And my vision, really is a future where all those models are in our platform, and all the different biases that come out of the models, we're using our blended approach to iron out some of those ethical biases that come with the fact that this model over here was trained on this set of data by this particular individual. So we also have a way that we can feed back from users. So best data is our users saying this wasn't offensive, or this was offensive, and feed that straight back so that our suppliers can kind of improve those models. So we think that positions is really, well, a) it means we can go and get new models. So we've had requests from a football club about could you do Thai and we went yes we can get you a Thai model. So we can support 23 languages, because we're standing on the shoulders of other academics that have already done this. Rather than doing that work ourselves. I mean, we're doing some research at the minute with the University of Bristol around content based bot detection, which is more unwanted content rather than hateful content. So where there are gaps in the market, you know, we'll either do it ourselves or we'll work with academic institutions to give them data because really what these people need is they need lots of data. And they can build new models and train them. So kind of our take is we sit in the middle, and we don't judge, we just know where the data has come from. We know how it's been trained. And then on the other side is, our customers that receive racism are the best judge of whether that content is hateful or not hateful. So we like to think that they always have the last word in deciding whether the model is right or not, you know, it's not up to us to be policing in the middle of that, really the end user and the more and more customers we get on board, the more and more data we get fed back, which means we'll get it right for more people more of the time.

Ben Whitelaw 15:44

Developing accurate ethical data models is only half of the battle though. I asked Matt, what the major platforms, apps and games can do to help get this technology in front of their users.

Matthew McGrory 15:57

I use the analogy to Microsoft having worked for Microsoft Partner previously, Microsoft has some core technology. And they created partner communities that were closer to end users and they were or they could ever be.

Ben Whitelaw 16:10

These partner communities are made up of certified companies that provide approved products in local markets.

Matthew McGrory 16:17

And I think Twitter, and Facebook and Instagram, and Tiktok, and time, are all creating access to their technology and their data in a similar way to create partner communities around them, that will fix niche problems in niche sectors in niche industries. And I think that's the only way that's gonna happen because Twitter are so big and Google are so big, and Facebook are so big, they're not going to be able to react to people in quicker fashion. As a kind of startup like we are and block party and body garden, these companies were able to react really quickly, regionalised in different areas of the world, different problems, because that's the other thing we found is what's hate speech in the UK is potentially different. Racism in South London versus the Bronx, versus Lagos, can be very different. And there's going to need to be regionalization for those types of those challenges. So I think it's the technology is at a point where it's affordable for startups to take that on, where it didn't used to be the problem is a big enough problem that people have gone enough is enough, let's really get this fixed. Now. I think the stance taken by the social media networks is that, you know, the internet is free and open. And it's a free space for free speech. And they're very sensitive to infringing on that free speech. And every time a new rule gets put in place, there's a danger each time that we stop things being said, that should be said, and it's like, you know, during my 18 months, I've read some of the philosophy I hadn't picked up for 15 years. But it kind of John Stuart Mill's concept of the harm principle.

Ben Whitelaw 18:03

For those not familiar with it, Mills's harm principle states that a person can do whatever he or she wants, as long as their actions do not harm others. But if they do, the society is able to prevent such actions taking place.

Matthew McGrory 18:16

The problem is, is the harm principle is a very subjective concept. And what's harmful to me will be completely different to what's harmful to an eight year old, to an old age pensioner to you, or depending on what country I'm in. So to make the network's put rules in place, across the board for everyone, I think it's just going to be a very challenging ask. So a really good example of that is like a banana emoji. If I send a banana emoji to a naturalist that's looking after monkeys, that's a positive thing I'm saying well done for looking after monkeys and keeping them safe. If I send that to a black footballer who's playing in the Premier League, that could very much be a racist connotation. And so we've got a real big grey area there to tackle and I think, centralising that, you know, for me, it's all about localization, and giving power to individuals to stop this happening for themselves and the ability to act on that where appropriate. So I think the reporting mechanisms need to be better. They need to be less litigious, less cumbersome. Focusing on that side of things, for me is much more effective than having a bureaucratic centre with a defined set of rules that apply to everyone.

Ben Whitelaw 19:40

Vidhya Ramalingam is founder and CEO of moonshot, a company working to not just remove hate speech and abuse when it's posted, but to stop it from happening in the first place.

Vidhya Ramalingam 19:51

we build technology that that serves two functions. So one is we work to try and build the evidence base. So really, some of the technology that we're doing is trying to gather data, analyse data and help us better understand the spread and the scale of this kind of content across social media platforms. And across the internet. Really, the second objective for our technology is actually to try and intervene. So we do an awful lot of work that is trying to directly interact with users that are posting hate content, posting extremist content, to try and offer them safer alternatives, and ultimately, to try and change their paths to steer them away from violence. And that's really the primary objective. That's our reason of being as an organisation. Yes, we do analysis, but the analysis is really to help inform those kinds of direct interventions with these with these audiences. Great.

Ben Whitelaw 20:40

So can you give us a specific example of one of those interventions that moonshot has used and the effect of that

Vidhya Ramalingam 20:48

one of the main methodologies that we've developed over the last few years is actually quite simple in its makeup, it involves the repurposing of advertising technology. So the same technology, the same ad tech that are used by big companies to try and get us to spend more money on their brands, we're using that sort of technology to try and reach audiences who through their online behaviours are indicating that they might be at risk of getting involved in violence. And what we're trying to do is use that ad tech to ensure if they're searching for hate content, searching for extremist content that we're placing advertisements to, to ensure were the very first option that appears for them. And so you know, a simple example of how this works is on Google. If you're searching on Google, in a country that we're running a project, if you're searching for, you know how to join a neo Nazi movement, you might then see an advertisement which was placed by us, which offers you a safer alternative, which connects you in with services connects you and maybe with an organisation that could try and work with you and steer you away from violence. So that's one method that we use. And it's something that actually we've scaled across the globe now.

Ben Whitelaw 21:55

Great. And what kinds of countries are you working in? And how do you decide where to do your work? Because there's an awful lot of challenges out there in the world right now related to the topics that moonshot covers.

Vidhya Ramalingam 22:09

Yeah, so we work globally, truly globally. I mean, we've done this work and tested these sorts of methods in over 30 countries at this stage. And the thing we find is, you know, if you're engaging with extremism, let's say you're engaging with Neo Nazism, in Australia, or in the United Kingdom, what it takes to engage with someone at risk of that form of violence is not hugely different from what it might take to engage with someone at risk of violence in, you know, in Sri Lanka, where Ultra nationalist

violence is becoming a problem or in Lebanon, if we're thinking about Salafi jihadism, the the underlying issues, and this really comes from our deep understanding of violent extremism, and what makes someone vulnerable. Ultimately, across the ideological spectrum, there's a whole range of human needs, and emotional needs that tend to be the drivers of violent extremism across the spectrum. And so for us, our approach has been, you know, let's let's build technology that helps us to reach these niche audiences online, in any part of the world, in any country. And then ultimately, the message that's going to be most effective with them, the services that we link them into, those are going to be locally specific, geographically specific, but oftentimes, we tend to see that even across geographical boundaries, the results are very similar. We tend to find, you know, if we're engaging with Salafi jihadist audiences in Indonesia, they are more likely to engage with content that addresses their emotional needs, or addresses some of their needs around employment, or, you know, and job seeking and some of the underlying drivers just like we find with Neo Nazis in America. So the results tend to be very similar, despite the geography.

Ben Whitelaw 23:50

Interesting. And let's talk a little bit about the kind of technology that you use to identify some of these patterns of behaviour. What signals? Do you kind of discern in people's online behaviour to be able to target them with specific messaging?

Vidhya Ramalingam 24:09

Well, people leave essentially what I would call a trail of clues behind in the online space. It can take the form of for example, keywords that you might enter into Google. If you're searching for hate content, it might take the form of posts and comments. So if you use racial slurs or hate language and public posts and comments, it might be that you engaged or liked or retweeted a piece of content that was shared by an influencer with an extremist movements. All of those activities are you know, they form a digital footprint that tells us that someone might be at risk. Now in order to do this work, we are not going into private conversations. We are not, you know, overstepping any boundaries when it comes to data privacy. This is the very same data that as I said earlier that companies are using to sell us products and to plays advertisements to get us to spend more money. We're just using these signals to try and ensure we can be there when someone who's at risk of violence needs us.

Ben Whitelaw 25:09

Yeah. And I presume that the success of your work is is somewhat difficult to measure or take some time to be able to validate. Can you talk a little bit about that about that process?

Vidhya Ramalingam 25:23

Yeah, the beauty of doing this work online is that there's actually a huge amount of data that we can work with to understand our impact. Now, the main challenge for us is not you know, were we able to reach were we able to get our content seen by someone? Were they able to watch it? Because that's the easy stuff. Those are the figures that, again, any brand agency or media company can access. For us the the crux of the problem is, can we actually measure long term change in individuals as a result of our engagement with them, and that's a challenge. And that'll continue to be a challenge for us as we test these methods, but to maybe give you a sense of some of the initial results that we've gotten over the years, maybe actually, one of the largest scale programmes that we ran was in the United

States surrounding the US presidential election around the time that it was very polarising over there. Lots of incitement of violence, lots of extremist groups mobilising. And we ran some pretty large scale campaigns to reach audiences that were engaging with white supremacist content with armed group content with election related conspiracy theories. And we tested all kinds of messages with those audiences, connected them in with services. And what we found was that the most engaging ad that we ran was an ad that said, that had the statement, anger and grief can be isolating. So on average, audiences were 17% more likely to click on that message than any others that we tested. And it got even more interesting if you drill down into very specific portions of that, very specific communities within that audience. So Q anon empathizers, they were most likely to engage with ads that said, Are you feeling angry? You're not alone, learn how to escape the anger and move forward. So, you know, again, in terms of long term change, it gets complex as to whether or not you can truly change behaviour. But we've had so many learnings over the years as a result of our evaluation efforts, just in terms of what messages are going to be most effective to start that conversation with someone who might need it.

Ben Whitelaw 27:22

Yeah, super interesting. Thanks for those examples. You've obviously made a huge amount of progress as far as extremism and violence and terrorism in particular, how much of what you've learned can be applied to other areas of you know, that constitute online harm, such as mis and disinformation, sensitive content, you know, hate speech and harassment at a lower level? Are they areas that you're looking to go into as well?

Vidhya Ramalingam 27:52

Yeah, it's a great question. Because when we set up moonshot, our first mission was to end violent extremism, my background was in the fight against white supremacist and far right terrorism and extremism. And extremism was the reason why we started the organisation. But over the years, as we started to test these methods across the globe, we started to realise that actually, the methods had a lot of application across other issues. So we now today and as of last year, we expanded our mission to cover online harms, our mission is now to end online harms, applying evidence, ethics and human rights, and what's contained within online harms. That includes everything from violent extremism and terrorism through to disinformation, myths, and disinformation, all the way through to child exploitation, trafficking, gender based violence, all kinds of harms that are perpetrated online that are facilitated through online engagement. And what we found is that the same methods that we developed to try and reach audiences who were indicating through their online footprint that they were getting involved in violent extremism, we can use those methods to try and identify perpetrators or potential perpetrators across those other issue areas. We've also found that the methods can apply to help us identify survivors, and identify victims and people who may be in need of support. So we have expanded our mission over the years. Obviously, the message that you that you send to those individuals will differ based on the harm. The service that we that we connect to them into will differ based on the harm, but the underlying technology is very similar and applies across the harm spectrum.

Ben Whitelaw 29:29

Great. That's really good to hear. One case that's been in the media recently is the case of Nina Patel, who might have read about who suffered harassment in the metaverse. How would moonshots

technology be used in that case? And what are your thoughts about kind of safety in the metaverse in general?

Vidhya Ramalingam 29:53

Yeah, um, what Nina experienced is, you know, is obviously horrifying and it should definitely raise alarm bells around how online harms are - how pervasive online harm is. Any new platform that is created, and this is really since the history of you know, since the beginning of the internet. You know, if there's a new online space created, humans will find a way to exploit it and to do nasty things on it. And so this is just a reality of the online world, the way that we work with the tech companies to try and solve this and resolve this. Yes, of course, we support moderation efforts. Because those are important, it's important to make sure that we are removing violent content hate content, such that it doesn't have the negative impact it can have on wider on the wider communities that are trying to safely enjoy that space. But the message I always tell the tech companies when we work with them is yes, you can remove the post, you can remove the comment, you can remove the video, but when you remove that post comment or video, the person who posted it still exists. And they may continue to come back to the platform, maybe under a new user name, they might continue to try and you know, try and push that behaviour, they may go on to another platform and try and try and push that same behaviour there. That person still exists and still poses a threat to us. And this is where you know, we have to remember that it's humans ultimately, that create these problems, not the technology itself. So this is where with the tech companies, we've really been working with them to implement what we call safety by design initiatives. So this is not just about moderation, but actually creating opportunities to use these platforms as as ways to reach people who might be at risk of violence and connect them in with help. So in the metaverse the metaverse is just like any other pot new platform new space online. The Metaverse needs to be thinking it needs to be implementing safety by design measures to ensure that they can be there when users need help.

Ben Whitelaw 31:53

Online violence, hate speech and harassment have been around as long as the web itself and don't feel like they're going anywhere, anytime soon. Is it realistic to think we can make any progress in the coming years? And will it get worse before it gets better?

Matthew McGrory 32:08

I think we're in the two to five year bracket of fixing this problem. I think, where does it go - imagine a world in five years time where we're all cleaned up, and everything's clean online. You know, this stuff still exists in society. And there's a societal problem to kick it, to sort this out. So from our perspective, in the football space, you can see just by the pieces they do on Match of the Day, organisations like kick it out who are really moved to being educational, going to schools, educating people about - this is a nice thing to say this isn't a nice thing to say - and showing people the impact of not saying nice things, and what that means in the real world and what that means to people's mental health. And hopefully, we'll move to a place where that education programme kind of kicks off in full flight coupled with the technology. And we get to a place in five years time where I suppose social media is like it was, you know, we talk about in the office here, it's like it was in 2007 2008, where you can have really robust debate, you felt really safe, no one was going to do a pile on on you because you had a specific point of view. Everyone was kind of quite respectful of that point of view. And hopefully those kinds of

communities will become more prevalent over the next kind of five years as we create more safe spaces for people to go.

Vidhya Ramalingam 33:30

Well, I'm actually hugely optimistic around what's possible. And I think oftentimes, when I have conversations with folks about this, they're surprised at my optimism given just the scale of awful content that we're we're looking at on a daily basis. The problems are huge, and the problems have certainly gotten worse. And actually, you know, I it takes a lot to shock me and actually, in the last 18 months, I have been shocked more than I care to admit around the scale of hate online. That said, you know, I am optimistic, in part because these same tools that are being abused and exploited for destructive purposes, those same tools can be repurposed for social good. I mean, I already mentioned the very simple methodology that we've designed around repurposing, advertising technology to connect people in with services that can save their lives and save other people's lives. That's just a small way in which this kind of technology can be used to really be there to help people when they're in a moment of need. I am really optimistic around you know, the possibilities for every tech platform to be using the power of their tools to connect people in with services when they need them, and when they're at risk of violence either experienced violence in themselves, or potentially perpetrating violence with others.

Ben Whitelaw 34:53

In the online world, as in the real world, violence, hate speech and harassment is not something that we can ever completely get rid of. Nina Jane Patel's experience, and that of millions of women and marginalised groups every single day are a telling reminder of that. However, safety tech companies are providing shoots of hope that one day, online abuse of all forms might just be an exception and not the norm. If you'd like to hear more about the way safety tech is tackling harassment and hate speech online, head to the safety tech innovation network, an international network dedicated to the promotion, collaboration, and industrial application of online safety technologies. Become a member to receive the latest information about safety tech events, and reminders about future episodes of the safety tech podcast. Thanks for joining me, and I'll see you next time. This has been a 4kicks production