

Safety Tech Podcast [Episode 1] – Safety Tech 101: an introduction to combatting online harms

Tue, 4/19 11:23AM • 35:07

SUMMARY KEYWORDS

safety, tech, companies, online, platforms, safer, technology, users, people, uk, harms, organisation, children, protecting, parents, content, providers, experience, term, toxicity

SPEAKERS

Sam Donaldson, Ben Whitelaw, Suki Fuller, Vicki Shotbolt

Ben Whitelaw 00:02

The web has changed everything we do. It's where we access information, do business, access, healthcare, use government services, shop, and even socialise. But while much of our regular daily activity has moved online, it's only relatively recently that the safeguards we'd expect to see in the real world have started to appear in the virtual one. You don't have to look hard to find stories of people being harassed, bullied, radicalised or misled online. That's where safety tech comes in. Safety Tech was coined in 2020, and is the umbrella term given to the tools and services that help platforms keep people safe online. It includes any technology that helps platforms, products and services, to block websites, moderate content that violates Terms of Service, highlight misinformation, or identify a user's age to prevent them from accessing adult only content. While it has potential safety tech is not without complexity. Critics say there is a risk of safety tech being too restrictive, relying too heavily on unaccountable technology and trying to solve the unsolvable. This is the safety tech podcast brought to you by the safety tech innovation network. In this series, we'll be exploring what it means to be safe online, and interrogating the future of safety tech, by speaking to academics, safety tech insiders, and the platforms that they work with. My name is Ben Whitelaw, founder and editor of everything in moderation, a weekly newsletter covering the very latest stories in online safety. In this first episode, we'll be talking to a host of people who can provide us with an introduction to this nascent burgeoning space. Thanks for joining us. Suki Fuller is a data lead storyteller, and strategic advisor for technology companies. Her background in strategic intelligence and startup coaching, has seen her take a keen interest in the safety tech sector. Over the last few years filla has been working with the Department of Digital Culture, Media and Sport to support the sector and introduce companies so the idea of creating safer online experiences. Yeah, Suki, how do you define safety tech.

Suki Fuller 02:31

So I always explain it to people as it's basically protecting the human experience in the online world, everything that you would expect in your physical presence being in there, I hate to use the term in real life, because it's all real life, but actually being physically present. And having that tangible feel of something is making sure that you can protect that same experience. When you're in the online world. Really it's protecting the human experience.

Ben Whitelaw 03:09

We have terms like brand safety, safety, tech trust and safety, cybersecurity, you know, how do you think about these concepts and the way that they interact each other in the work that you do?

Suki Fuller 03:23

The way that I see it, it's just like cybersecurity is the protection of your files, your data, you know, that sort of ambiguous, because they're not quite tangible. They're not. They're not human, tangible. And safety tech is the human side, those two are two different sort of parallel, and where they should be sitting at the top. Everything else Trust and Safety brands, whatever you want to call it this week. I really do you think they should come under safety tech? I think safety tech needs to be at the top and everything else needs to be under that umbrella.

Ben Whitelaw 04:02

How realistic is it that we can maintain the human experience of real life of our physical selves online?

Suki Fuller 04:12

I don't know if it's realistic is the word I would say. Is it likely? Or is it highly likely? How much of it can you protect? Because we already know we can't protect the human experience. When we are actually in the physical world. We are not able to rest we wouldn't have murders we wouldn't have, you know all of these crimes that happen we can only protect it to a certain degree. So it's a matter of how much of that experience you can protect in the online world. And I would say you can protect I would say most of it. I'd say the emotional harm is the part where safety tech is probably going to be the most beneficial. The emotional, the psychological traumas, the psychological benefit. Just the emotional benefits of what safety tech will probably protect most physically, we're gonna get into that with, you know, VR and XR and MR. The whole metaverse. That's a whole nother layer. And we've only just gotten to the point where we're talking about safety tech and web 2.0. And now we've added in three, what are we going to do, we have to catch up, and it needs to be pretty darn fast.

Ben Whitelaw 05:26

Do you think the average person understands that the role of safety tech isn't to reduce harm to zero?

Suki Fuller 05:34

Most people when they hear of anything that has to do with safety, cybersecurity, or any aspect of law enforcement, they believe that it should be zero. There's a sort of a zero tolerance mindset, when it comes down to how much you can be protected. Just like with your regular law enforcement, your police officer on the street, they just think there should be zero crime. Because you have police. And that's just not the that's not human nature. That's not the way of the natural law of the world of human existence. And I always tell people, safety tech is a way to reduce the possibilities. But at the same

time, while we're reducing the possibilities, we're able to actually now identify more of those possibilities. Whereas before, we had no measure whatsoever, it was just a free for all it was like the Wild West.

Ben Whitelaw 06:30

Yeah, interesting. What do you think are the competitive advantages for companies that get safety tech?

Suki Fuller 06:39

I always look at it like this, you have your companies that are new. And I always say, Well, if you don't integrate this now, at the very beginning of your product journey - I was actually having a conversation with some people when they were talking about NFTs. And they had quite a successful range. And I said, well, where's your safety in this? The possibility exists for your average person that is selling something like an NFT, that someone can get some malicious content in there. So why would you not build those aspects into your product at the very beginning? Why would you not make it safer? One is better financially fiscally, for your company, is protecting your ROI. And two is protecting your user because that's protecting your reputation as a brand. Three is protecting your brand identity, because someone's not messing with it. So I asked, why would you not do that? And they said, Oh, you know, we're gonna build that in later, later is too late. You know, it's one of the most frustrating things, especially as an intelligence analyst is to say, oh, we'll just wait, you know, something, we'll see if something happens. Why do you want to do that you mitigate the risk? You know, you don't go outside without an umbrella going, Oh, I'll wait until it rains. And then I'll go get an umbrella. No you take the umbrella when you're leaving a house because the possibility exists? Why would you not do that is so frustrating. And I think that is the definite advantage that companies that are actually thinking about safety tech, the companies that are doing safety tech that they have when they go to others and say you need to do this now, because this is not an issue for you today, but it will likely be an issue for you tomorrow morning. And it may already be an issue, you just may not be aware of it. And a lot of people are not thinking about that as we go into web three. It is going to become a bigger issue.

Ben Whitelaw 08:46

Yeah, it's really interesting. I love that umbrella analogy. What companies are you really excited about in the safety tech space? Which ones are you looking out for?

Suki Fuller 08:55

I can't say that. These are my people. I am really impressed with Yobu. And the fact that on their website, they had their nice little pop up and this is a social media website for teens. And they did this first and I was like of course they did because they were thinking about the protection era and it was that whole you know, are you sure you want to post this? It was basically think twice and then you know now Twitter has embedded that I think it was last year they embedded but I really like what you bill has done there because it started a nice little seed and it put everybody else on notice like if we can do it and we're a small company definitely can do it. I like the age estimation that Yoti's doing I really am very and they know I have I have so much love for them. No offence everybody else but I do. I think it's just going for, are you legal? Are you allowed to be doing this thing online? Or this thing in the

supermarket? Are you really supposed to be able to buy alcohol? I like what they're doing there. And I really am interested to see how far they can take that.

Ben Whitelaw 10:20

One man that is familiar with these companies is Sam Donaldson. Sam is the director of perspective economics. And in 2020, he wrote the first report of its kind on the safety tech sector in the UK, on behalf of the Department of Digital Culture, Media and sports. He found that in the UK alone, there were 70 different safety tech providers, and has since seen the number of companies in the space and their revenue grow quickly. Sam, if you could explain for the listeners, what are the findings of the research that you've done around safety tech recently?

Sam Donaldson 10:57

Yeah, thanks. Thanks, Ben. So we kind of first came in line with this idea of of safety tech relief. In 2019, we started a piece of research with the department for Digital Culture, Media and sports in the UK, it was all really about kind of seeing on the ground that there were a number of companies that looked a little bit different to what we might typically consider the cybersecurity or just general safety. And DCMS had this sort of hypothesis that there appeared to be the sort of the seedlings of an emerging market related to online safety technology. We have previously worked with the CMS looking at the size and scale of the UK cybersecurity industry. And even when doing that research, we actually started to find these companies that were kind of on the fringes that we sort of said, well, they're not they're not necessarily cybersecurity companies, but they're mainly focused on protecting individuals, typically children, or typically focused on areas like network filtering, that were sort of we've known for many years. So we've really started to find something kind of distinct coming through initially in the research. When we started that research, you know, I admit at the time I kind of, I had no idea really how many companies we're going to find that sort of met this, this criteria met this definition. So the first stage was really about, what is our definition of, of safety technology, but how broad or how tight do we want to make that? And we came up with a really broad definition that - a safety tech provider really is anyone who's developing technology or a solution to try to help either facilitate an online experience that is safer for the user. Or is protecting users from harmful or illegal content or conduct. And within that definition, then we kind of started to break that down a little bit. So what what does that actually include? One thing that we were very struck with really is the breadth of technologies that exist to try to keep users away from online harms. So what we find initially in that 2020 Study was 70 providers in the UK offering safety technology, that sense grown to over 100. And I suspect there are many more, we've also find over 500 companies that kind of fit into this category, globally as well.

Ben Whitelaw 13:10

So can you kind of tease apart what some of these companies do for us then?

Sam Donaldson 13:15

Yeah, absolutely. We categorise them into five key areas that we think are particularly relevant there can be well understood by by industry, and by end users, we start off with an area kind of called system wide governance. And in short, really, what that's about is about tracing and removing known illegal content, right, we also know what that is. So it could be child sexual abuse material, it could be

terrorism. So again, a couple of examples, you know, the Internet Watch foundation come to mind QModo, in London, Cyan forensics, Scotland as well.

Ben Whitelaw 13:49

Just a note to say, since recording this conversation, signed forensics have changed its name to Cyacomb limited. Back to Sam.

Sam Donaldson 13:57

anybody kind of issuing those technologies to help kind of get that most harmful illegal content away from end users and off the internet in general, and also hold people committing those types of activities to account. And then the second one is about the idea of platform level factors. So what are platforms doing? What can platforms do? What technologies can they use to keep their users safer online. And those platforms again, have two real responsibilities. One is on illegal content, as I talked about making sure that is not on their platforms, making sure that there are sufficient identifiers to be able to remove that. But actually, then you end up with this sort of the term of the lawful but awful, the kind of harmful material that is not necessarily in its own right, illegal, but it's very problematic for platforms to contain. So ideas of sort of content moderation, picking up particularly hurtful terms that might be in text or image or video. I'll give a couple of examples of companies in that category. Crisp based in Leeds come to mind, Spirit AI, the moderation company in London, ultimately trying to sort of remove toxicity keep platform safer, cleaner and just more enjoyable for users to be on. And then the other area of of the sort of the platform level that sometimes gets a little bit overlooked. But again, it's a really important part of the puzzle is this idea of ah assurance, it's appropriate design. The question for platforms bringing in this sort of technology is, do you know the age of your user? Have you verified your user? Have you made sure that your platform is actually safe for not only under 18s, but also the certain consents in place for under 13s as well. So that's the platform level, obviously, the endpoints and this is kind of cybersecurity term I guess, but thinking about the device, the laptop, the phone that a particular user has, typically a child that could be anyone, this idea of endpoint protection, classified this into two sort of subcategories. So the first is user protection, not necessarily be conflated with parental controls, which I think is sometimes a slightly outdated term to use. But what is it that can be put on to say, a child's device to make sure that it is safe out of the box. So again, a good example of that - a company called Safe2net, who built in automatic filtering, and also automatic behavioural checks on users as well. And then network filtering, is that last category. So again, whilst it's been around for a sort of 30 plus years, there are a number of really established filtering companies in the UK that do a really good job at identifying content, but particularly things like users and schools, children should be accessing. And then lastly, and by absolutely no means least, this idea of information governance, which is a really unsexy term for a very interesting concept. But essentially, that's about looking at how can technology be used to detect disinformation and misinformation in the public domain, lots of really interesting companies in that space, but it's on the fact checking side, even seen a move into content provenance as well, and ultimately, calling out this information and calling out the quality of particular sources that are being used online. So companies like full facts, use guides that come to mind that matter, logically, quite a few we've seen particularly sensitive the 2016 presidential elections coming through. So hopefully, that again, I did say it's very broad for sort of over 100 providers, and there's a lot to that, but hopefully that gives you a flavour of what's kind of on the market.

Ben Whitelaw 17:15

What's very striking, I think, from the report is the growth of this sector. What do you think is behind that?

Sam Donaldson 17:25

So just to give a little bit of context as to those figures, when we started the very first safety tech, sectoral analysis exercise, I mentioned that there were 70 providers, the revenues, were just just over 200 million, which again, for such a small number of companies is really particularly strong. And then we came back to it last year, we find that the revenues have increased from 200 million to over 300 million, essentially, when Wow. So what we're looking at at the moment, and what we've seen in the last five years is a compound annual growth rate of 30 to 40%. year on year on year. And a lot of them are not actually selling directly to concern parents are users, they're selling their services to other businesses. So they're saying, we've developed a number of in house bespoke algorithms for content moderation, that can work for your company, we can work with you. And we can identify a problem that is costing you money as a company. And I think that driver of safety tech being used to make the user experience more positive, and in turn actually created benefits for the company buying it has been an incredibly important role in all of this. The second, and it's not even the elephant in the room really, but regulation has happened and is coming. And it's a mix of or somewhere in between the midst of both of those at the minute. But actually, if you look at the interpretation of the GDPR, and I mentioned the ICO age appropriate design code as well, many large online platforms already have a number of requirements to implement for user safety user controls, as is, even things like some of the really big growth areas here we see particularly strong one is on the gaming side, online gaming has had its real issues with toxicity and problematic behaviour for many years. And I remember seeing a number of providers come to the market and say we can actually help you resolve toxicity in the markets. As the technology has improved. I think it's no longer a case of saying, you know, there's some key words we can remove from the platform. It's actually about developing bespoke models of moderation that I think some of the larger gaming companies, media outlets, social media platforms are starting to wake up to, and it's still a very small part of a really big market. So I think there's still a lot more growth to come.

Ben Whitelaw 19:33

Yeah. And I want to ask you about the UK is position in the kind of global market. Now, obviously, the government have ambitions to be the world's leader in safety tech. Can you give us a sense of how they're doing as far as the aim goes, and who which other countries are also, you know, very advanced in in their safety tech sector.

Sam Donaldson 19:59

So I think one of the points of leadership I think the UK Government have shown with respect to safety tech is, is the fact that this research was very much it's kind of the first of its kind that we had seen globally to try to understand the shape of this market. And we did do a lot of research trying to benchmark the UK's position against global provision. So I mentioned earlier that we've identified just over 500 providers of what we think kind of match the definition globally, round programme that would suggest that one in four or one in five companies globally are sort of in the UK, we've seen, countries like the US and Israel, have been particularly strong in developing a number of very large migration companies, you also see quite a few interesting examples across Western Europe as well. So

Germany, there are a number of interesting companies coming through following some of their legislative push there on the toxicity of hate speech, Denmark, Sweden, Norway, again, some really interesting examples around there. But I think one of the really interesting things the UK Government has done, particularly with the safety tech innovation network is opening that up globally. I think that's a really, really important narrative to do. Because ultimately, it can't just be about the UK tackling these problems. The other interesting kind of nugget of information as well, it's in the report is that I think over half of the companies we identified in the state export to international markets. So we're developing a lot of in house capability in the UK, and a lot of expertise and data science and ethics, and all of the things that come with these technologies. But we're selling these globally, right across the world. And I think so long as the UK continues to have an international focus that should continue to put it in good stead as a global leader in safety tech.

Ben Whitelaw 21:44

Among the groups that spend a lot of time online, and that's the most vulnerable to online harms are children and young people. Vicki Shotbolt is the founder and CEO of parent zone, which works with technology companies and third sector organisations to help them create child friendly products and experiences. She has been working in the family sector for over 20 years. And he's on the executive board of the UK Council of child Internet security. So she knows the risks that children face online.

Vicki Shotbolt 22:14

It's difficult to capture the essence of what parents zone is all about to be completely honest. But fundamentally, we support families. And that includes children, parents, and the professionals that work with them to try to navigate the messy complexity of growing up in a digital world. And that takes so many forms. It can be face to face work, it can be writing and designing education programmes, it's it's all kinds of stuff. But with that core thread of trying to help kids get the most out of a digital growing up really

Ben Whitelaw 22:48

what attracted you to the safety tech sector in the first place?

Vicki Shotbolt 22:52

That's a great question. And I don't know that I was attracted to the safety tech sector. if I'm completely honest, my background is working for children's charity, I used to work the Children's Society after I worked for the Children's Society, I worked for the National Family and parenting Institute. And the mission there was to try and make Britain a little bit more family friendly. We were sitting right at the very bottom of the league tables in terms of how family friendly we were as a country. And I suppose I moved into doing the work that we do at parents home, because the world was changing so much. I remember doing some work with a mobile phone company. And they said, you know, really soon, kids are going to be accessing the internet on their mobile phone. And I was like, That's ridiculous. Nobody's gonna do that. We're talking the days of the David Beckham, flip phone. No, ain't gonna happen. And so my interest is really much more in tech as a transformative force in a child's life. safety tech kind of comes in as a, how can we use technology to try and make sure that that experience is a positive one, and one that is not risk free? Because you can't eliminate risk from a child's life, but as harm free as we can possibly make it?

Ben Whitelaw 24:06

And, you know, do you have any examples of specific kind of children or families that you've worked with? And can you kind of explain a bit about the effects of harms when it happens? Because I don't know if many people understand that unless it's happened to themselves.

Vicki Shotbolt 24:23

Oh, there are so many examples that stick in my head. I mean, the I can think of a parent that we worked with who had a relationship online, and it's the kind of classic nightmare scenario that a parent goes through where the young person was talking to somebody who didn't have that young person's best interests at heart and it became a toxic relationship that tipped over into something that ended up requiring the police to be involved. And there's nothing more terrifying than knowing that your child is being harmed online and it feels so out of control. It's not like somebody in the state read that at least you can look at me, you know, grab your child's hand and walk them away from that person, it feels really invisible, really, really frightening. And the effect lasts for years, you know, children dropping out of school because they're too frightened to go to school or they're too depressed and upset to go to school, or they have to go through the pain of a court case, or whatever it might be. They are really horrific real life harms to children, for some things that I've started online. So when I think about harms to children, I think about that massive spectrum, from a child who is being bullied online, or maybe experiencing overuse and the early signs of something like a gaming addiction, or, you know, just finding it difficult to connect online, you know, feeling excluded from your social group, all of those things can be experienced as harm all the way through to the other end of the spectrum where a child might be being radicalised might be ending up in criminal behaviours online, or might be being sexually exploited. So, you know, harm is a massive, massive span. And I say, again, it's different to risk, you know, children experience, that same range of risks all the time. Are the problems that the internet significant and large, and do they impact millions of people? Yes, they do.

Ben Whitelaw 26:18

What role do you think that safety tech has in keeping those people safe?

Vicki Shotbolt 26:23

As an organisation, we are really interested in helping children to thrive thinking about the personal attributes that a child needs in order to become digitally resilient. So the role that it plays is much more about reinforcing safe boundaries, it is, you know, is the soft landing spot underneath a slide it's it's the safety lock on a on a bottle of tablets, it's, it's the thing that will put in some measure of protection in the context of all of the other much more important stuff in some ways, which is the parenting, the education, the child's personal attributes.

Ben Whitelaw 27:02

Can you talk a little about your your growth as a company, as the safety tech sector has grown, and perhaps some of the things that you've noticed over the last five years in particular,

Vicki Shotbolt 27:14

we do do an awful lot of work around safety tech, and with organisations that are developing and creating safety tech. And I think what we observe is that it's becoming embedded into many more platforms, many more organisations are including it from the get go, you know, from that moment, they're thinking about building their platform and element of it's going to be around safety tech. And for us, we are thinking about creative ways that we can build tech that will increase safety. So we're doing a project at the moment with Benardos, for example, which is all about supporting social workers to have the technology they need to answer in the moment questions. So it's a sector that's growing massively, and really fast, we've grown from an organisation that had, you know, me and the dog to there now being 30 of us and still the dog. So we've grown really, really quite quickly because you know, the spaces just massively grown.

Ben Whitelaw 28:14

Glad to hear about the dog still being part of the team

Vicki Shotbolt 28:16

The dog is a critical part of the team.

Ben Whitelaw 28:20

Can you talk about those companies, who they are, and what opportunities they are probably missing, as far as you're concerned, to create experiences that are safer for children in particular?

Vicki Shotbolt 28:33

Well I suppose the sort of companies that we work with, tend to be the really big guys. So you know, we work with organisations like Google and Facebook, and what is and I think people don't always appreciate how much effort they put into safety tech, you know, we have this kind of public perception of them as platforms that are all about the bottom line. And they don't worry about things like making their spaces safe. In actual fact, they have hundreds of thousands of engineers focused on building safer products. And I guess where we work with them is talk to them about what the ways are to build pro social experiences to give tools to parents, like family link where parents are in control, and all they've done is use clever engineering to build something that gives parents the tools that they need. So you know, I see loads of work going on. And I think where I see when I look at the safety tech sector, especially emerging companies, focusing very much on that protection lockdown limit end of the spectrum, and maybe not enough on the - how can we facilitate and enable and create safe places, safer places that aren't locked down that don't prohibit a child's access? And I think that's the exciting sweet spot. But I think safety tech is at its best when you don't even know that it's there, then it's really working.

Ben Whitelaw 30:07

The topic of online harms has received a lot of attention over the last few years. And internet platforms and services in particular, have faced increased scrutiny about harms that people are exposed to when online. Governments around the world have responded in kind by proposing or enacting legislation that makes companies liable for harmful content. The European Commission's Digital Services Act and digital markets Act, as well as the online safety bill here in the UK, are just a few examples of proposals that we can expect to see become law in the course of 2022. Whether they will make the internet a

safer place, we'll have to wait and see. With so much happening, I wondered where our guests that will be in five years from now. What can companies working in this space expect? And how will that affect people like you and me, as we move around the web, and live online? Suki paint me a picture of what the web looks like, in five years time, in regards to some of these safety tech issues that we've seen.

Suki Fuller 31:16

The one area that I see that we definitely need to work on is really the bottom of the bottom, people that don't have access to the web, they're going to get left out of this, already they're being left out. The people that don't have money that don't have access to the web that don't have access to faster, you know, they don't have, they're not going to have 5G, whatever, they barely have 2G, you know, they don't even have working phone lines, they don't have the infrastructure. So those countries that don't have the infrastructure, they're gonna get left. And they don't have the ability to just hop, skip and jump to where we are, we need to make sure that it's not cost prohibitive for people to be doing anything in safety tech. And that is just the norm. It's just something that when you start a company, you automatically integrate that, just like people do in Lean and Agile, it becomes part of that cycle. That's what I would love to see

Sam Donaldson 32:14

what I would like to see and what I really would be keen to see. And I think this will start to kind of come through in the figures over the next few years, is just to see that increased adoption of safety technology across a real breadth of sectors. Any organisation that has an online presence, really, once you hit a certain size, you should start to consider having a trust and safety team in place. What I would like to see over the next few years is kind of a thoughtful sense of leadership as to - here are all of the technologies here are all the approaches that exist online. Which of those can I as a business or as an organisation implement in an easy to use way that is going to maximise the experience for our users, and ultimately improves the safety of the users. Sometimes Safety Tech, I think get conflated with just - oh just keep the harmful material off or just keep my users not swearing at each other and giving each other abuse. But actually, there's so much more to it than that. It's how do we make sure that hate groups something inadvertently funded through your platform? How do you make sure that you're not accidentally sharing this information? How do you ensure that your users are actually just having a good experience on your platform, and you're retaining them? So there are so many more questions that can be asked. But I think as long as there's a thoughtful leadership, the ask those questions, I hope is that the internet should be in a better place.

Vicki Shotbolt 33:30

So I'm really excited by the idea that that it will become integral to tech. So it won't be sort of, you know, you'll get your you get your tech thing, whatever it might be a game or a platform or an app or whatever it might be. And then you go when you have to find the bolt on bit of safety tech that's going to make the experience better. I'm really hopeful that in five years time, what will have a little bit like Pay Pal made the experience of shopping something online buying something online really seamless. My hope is going to be much more like that so it becomes an integrated part of the experience.

Ben Whitelaw 34:08

Over the next five episodes, we'll be speaking to people that have had a real life experience of online harms, as well as representatives of safety tech companies that are trying to create safer digital spaces. We'll hear more about some of the problems we've touched on in this episode, including misinformation, harassment, and hate speech, and try to understand whether an internet without harm is possible, or even something we should be aiming for. For more information on all things safety tech, 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. This has been a 4kicks productions podcast