

Safety Tech Podcast [Episode 6] – The power of design: how subtle choices can prevent users experiencing risk and harm

Wed, 5/25 10:25AM • 40:20

SUMMARY KEYWORDS

grindr, users, people, safety, design, privacy, platform, online, app, children, designers, services, thinking, world, product, important, code, messages, harm, choices

SPEAKERS

Iain Drennan, Georgina Bourke, Ben Whitelaw, Alice Hunsberger

Ben Whitelaw

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, in the final episode of this series, we're spotlighting one possible solution to the online harms crisis, safety by design. This new approach to user safety, which emphasises harm prevention at all stages of the design and development process has the potential to revolutionise how people, especially children, experience online products and services. But what does it take to make digital spaces safer from the get go? Thanks for joining us.

Alice Hunsberger

Some are using it to find connections in real life. And some just find connections online.

Ben Whitelaw

This is Alice Hunsberger, the senior director of customer experience at Grindr, an app designed to help LGBTQ plus people find and meet each other

Alice Hunsberger

dating and finding connection is a emotional thing. People feel lonely and are trying to find community and connection. And so sometimes feelings get heightened. And there's you know, harassment and bullying and discrimination and you know, not great behaviour that sort of stems from that rejection. And then, you know, we also see, as with any platform with millions of users, people who are trying to take advantage of them,

Ben Whitelaw

Grindr has used safety by design principles to good effect over the last few years, to be inclusive of the specific challenges that LGBTQ plus people face when meeting others online. I wanted to find out how Alice and her team proactively addressed certain types of harm. But before we got into that, I asked them what the most prevalent issues were that affect grinders 13 million plus users.

Alice Hunsberger

So you see a little bit of everything. But right now, we're really focusing on spam, that sort of the biggest volume, the highest number of requests that we see, you know, sort of no surprise, because there are entire industries of people whose nine to five job is to try to get through all our defences. So sort of an adversarial problem, we, we put in one fix, and they're super smart and come up with a workaround. So that's sort of a constant, ongoing battle that we have in every other platform.

Ben Whitelaw

Maybe we can dig into that. And so what are the kinds of fixes that you would put in place to try and combat something like spam? Can you take us through those? We have

Alice Hunsberger

machine learning and artificial intelligence systems that check upon signup, they check again later, we have, you know, all sorts of smart rules and workflows for, you know, unusual behaviour. Most of these accounts are created by computer programmes. And so we have automatic systems to try to detect the anomalies. One thing, you know, that I think is interesting, is we always tried to sort of balance efficiency, but especially efficiency and scale something like spam, with also making sure that we don't cause any issues or harm to our users. So you know, something specific that Grindr does that not all platforms do is there are a lot of trans and non binary users on Grindr, and they have every right to be there just like anybody else. But unfortunately, there are some people who don't think that and so trans non binary users will be flagged recorded by other users disproportionately highly often as fake or spam accounts. Meaning I think this person's identity is fake, not that it's an actual scammer. And so we don't automatically ban any user based on another user's report of spam or fake activity. And you might think, Oh, if an account gets 10 flags from other users all saying that it's a fake account, obviously, it must be so we'll just automatically ban it and that's the best thing for the community. But we sort of take extra care, especially because we, you know, obviously that bias and we don't want to increase it with our own moderation systems.

Ben Whitelaw

And that's obviously a really traumatic thing for that account to have to do if it was taken down with no explanation, right, I can totally see why you do that. What happens at that point in the process, if an account in that example you gave has been flagged, who is the person or who is the team that is, is triaging, though,

Alice Hunsberger

we have a content moderation team, of humans, who look at it, they're all sort of trained specifically, not just, you know, for our community guidelines and rules, but they also go through bias training, they go through specific training on the LGBTQ plus community so that they understand, you know, specifics to the community. And so, you know, they go through and review these accounts manually.

Ben Whitelaw

What can be done to mitigate some of the harms that crop up in a platform in advance? Are there ways that Grindr is trying to prevent these things happening before they come to the fore? Yeah, of course,

Alice Hunsberger

being proactive is hugely important. And Grindr considers safety by design, sort of from the beginning of every project that we do. So some lifespan is always going to be sort of reactive, and we're gonna go back and forth. But there's also, you know, with any new product, we consider, what are the implications for moderation? What are the ways that it could be exploited? What are the ways that we can really design it and make it great from the beginning? So you know, one good example that we just rolled out globally for everybody at Grindr is albums photo albums that you can send to any other user on the platform. And I worked really closely with the product team to sort of consider like, often, we assume these are going to be incredibly personal images, that people like to send each other on dating apps. And so really limiting who in the company is allowed to even see these images to begin with, right to make sure that user privacy is respected the whole way. But then also, what is the flow for users to send these albums to each other. So we designed consent into the album sharing. So the whole album is blurred, you don't see what the picture is, you click to accept it. If you don't want to see it, you can remove the album on either and and so that's sort of a really nice way that we sort of considered not only how to deal with these albums, if any of the content does get reported, but like, how do we prevent, you know, issues to begin with, through this sort of consent design?

Ben Whitelaw

One of the things that I think other platforms and services and apps have maybe not thought enough about is how a service can be used for nefarious means, right? What's the worst case scenario that can happen? How do you bake that into some of your product work?

Alice Hunsberger

I would love for us to live in a world where the being gay or trans or you know, just part of the LGBTQ community is normal and accepted. And, you know, not any kind of issue or problem. But unfortunately, you know, there are some countries where it's still illegal to be gay, and lots of places where people face discrimination. And so our user base is especially sensitive. And, you know, so we really, I think, have a heightened sense of the issues that our users could face. You know, one example of thinking about how our users can come across problems is like, just the app itself, right? Like, it would be wonderful if we lived in a world where having Grindr on your phone was like a non issue. But for some people, even just having an app icon on their phone, can potentially put them into awkward situations at the mildest or dangerous situations. So we have alternate app icons that people can use so that the app you don't see the very famous mask logo, you see it note pad or clock or something like that. We also have a PIN code so that in order to even open the app, you need to put in a pin.

Ben Whitelaw

Yeah, that's really interesting. I was really interested in some of the features that Grindr released over the last few years like the ability to unsend messages and also block the taking of screenshots. How

have they had an effect On harms within the community and, you know, how do you measure the effects of those kinds of features?

Alice Hunsberger

Those features, many of them were developed thinking about these areas in the world where it is more dangerous or risky to have Grindr use Grindr. So we sort of have this classification system that our director of Grindr for quality, which is sort of like our activist, social justice, global perspective arm of the company, he keeps an eye on which areas of the world are progressive and relatively safe and which ones aren't and then talks to people in those communities, and says, What do you need to feel more more safe? And so we have sort of extended suite of features that we give for free to everybody who's in those areas, some of the ones you mentioned are included, we also do extra safety messaging. And then you know, some of them will give even to people who are us or anywhere, because many of them are useful for everybody. It's always tricky, because you asked about measurement, and some of these things are just the right thing to do. And so we're not always looking at metrics for user engagement to say, Okay, we put out this feature, and now more people are messaging each other because we allow the messages to disappear after a certain amount of time or whatever. It's sort of like, we have a general philosophy, that's what good is, what is good for our users is good for us. And developing features that help people feel in control of their account and their data and their privacy and what they share and what they don't share is good for them. And so it's good for us as well.

Ben Whitelaw

You mentioned the kind of safety messages there. Can you talk about the safety messages that Grindr uses, how they appear in the app, what they're used for?

Alice Hunsberger

Yeah, so one of the first projects that I took on when I started at Grindr, I was rewriting and revising our safety tips. And they were already in a good place. But you know, there's always more details to add. We haven't translated into tool into your 30 languages, and they're available for anybody to see in our help centre. But people who are in areas where there's you know, more of a risk, then we send inbox message inside the app, reminding people of the you know, safety tips or things like, be careful before you meet somebody, do some research, see if you can see it social media link or something else about somebody for you meet up in a public place. So we'll send it inside the app. And you know, inside the inbox, because again, sending like a push notification, or an email outside of the app to somebody who's disguising the app as like a notebook app, because you know, it's risky for them isn't a great idea. So we do it all inside of the app, and you know, sort of have a different cadence for how often we send it out. So in some places, we send them every day, some places, globally, we send them out every six months. And it's always available on the app.

Ben Whitelaw

Let's admit, however, that there is sometimes a tension between keeping people safe online, protecting their privacy, and maintaining their right to speech.

Alice Hunsberger

If I put my head of trust and safety hat on, I want all of the information about every single user forever. I want all of their like government IDs, I want their AI every single IP address going back for all time, any tiny bit of personal data, because that is what helps you detect bad actors. And it's what helps you keep them off the platform. So from a safety point of view, all the personal data is a good thing. From a privacy point of view, obviously, you want to protect that personal data and people have the right to be synonymous on the internet and not have their whole identity and especially on Grindr, where it's illegal to be gay in many countries, linking your government ID and your real world identity when you know with a calm money that is associated with the gay community could land you in jail if that data ever gets out, which, of course, we would never let happen. But it is obviously something that people worry about. So there's this tension on any platform. But you know, especially because of the sensitive issues with Grindr with how much information do we ask from users, what do we store? How long do we store it for? And you know, what is that like perfect balance between privacy and safety. And, you know, we get hit from both sides saying, you're not keeping your user safe enough. But also, you're asking for too much information. And you can't have it both ways, like, there has to be a compromise, and you have to meet somewhere in the middle. And if we had a perfectly private app, it wouldn't be super safe. And if we had a super safe app, it wouldn't be super private. And so that's sort of an interesting tension there.

Ben Whitelaw

Rather than make the decision for them, Alice's upsides and giving users control over their choices, and designing control panels and settings, where they can select and change their own preferences.

Alice Hunsberger

Some users will be like, I don't want my messages to be read by anybody else. I don't want machine learning tagging my messages. I don't want anybody else's messages to be censored by some kind of random algorithm. I want complete free speech and expression and also privacy. And you have other people who are like, I am very triggered by this particular type of category, it's really harmful to my mental health to ever see anybody talk about this thing. And it's as a platform, it's your obligation to make sure that I never have to see it, which then means the platform has to be monitoring all the messages and censoring things and removing free speech and, you know, digging in there. And so the answer has to be giving users control over that. But also, you know, then you sort of turn into having this like, suite of very granular options for people. And it sounds great on paper. But we all know that users of apps generally want like a very intuitive experience, and they don't want to spend 30 minutes setting toggles about like, they don't think about privacy and safety and free speech and like where they are in that triangle the same way that I do, it's not their job, why should they have to spend time doing it. And so they also want ease of use, and for platforms to kind of figure it out for them to some extent. And I think that it's something that the industry just hasn't completely figured out yet. We're still really in the early stages of testing what people's appetite is for these kinds of specific controls and what the balance is, you know, in the regulatory world, as well as just policy for the balance between privacy and safety. And so it's very difficult.

Ben Whitelaw

Someone who's thinking a great deal about the trade off between privacy, safety and speech is Georgina Burke, a specialist in user experience design, and the principal technology advisor at the ICO,

Georgina Bourke

the ICO is the Information Commissioner's Office, that's the UK Data Protection regulator, I have this really exciting opportunity to help translate data protection, policy and law, it's often quite legal into practical tools and guidance for designers and product teams that they can understand and use. So that's the area that I focus on.

Ben Whitelaw

The ICO's job is to promote good practice in handling personal data. And in 2020, it formally launched the children's code, also known as the age appropriate design code to oversee the way that children's online data is stored, used and managed.

Georgina Bourke

The Children's code is a code of practice that's designed to protect children online. There are 15 standards that look at different aspects of the way that you can handle a children's data, the way that it should be used and the way that services should be designed for different age children kind of risks that you should be thinking about. The code is quite broad in scope. So it really covers any services likely to be accessed by children. So it doesn't even need to be, you know, explicitly designed for children. But if you know that children of different ages are using your service, then you need to design your service to meet their needs. The code has come about through a number of harms that we're seeing online and the need for regulators to set out how people should be using children's data and protecting children. And really seeks to include children in the digital world. So the code doesn't want to exclude children from it or say that you shouldn't be processing children's data or children's search shouldn't be using the services, but really create services that are safe for everyone to use.

Ben Whitelaw

Online Services likely to be accessed by children have had 12 months to comply with the code. Meaning we've recently started seeing the effects of the new standards,

Georgina Bourke

it's had a really positive impact so far. So it came into force in September 2021. And we've already seen some really positive changes from some of the big tech companies. So things like changing targeted ads and personalization for children, making sure that they're switched off, making sure that settings are high privacy by default for children, and other safety features that people might not have considered before. So switching notifications, off bedtime, things like that. It's already having a ripple effect, not just in the UK, but on the global scale as well. And we're starting to see other regulators in other countries, using it as a blueprint almost to look at their own laws and regulations. So hopefully, it will influence other countries and other codes of practice, that will just lead to stronger privacy protection for for children, you know, all all over the world.

Ben Whitelaw

Lots of people won't be aware of the subtleties of the apps and digital services they use. So I wondered if you could explain a little bit about what we mean, when we talk about design choices.

Georgina Bourke

You're absolutely right, that people don't often think about the design, and the decisions that happen in the background when you're using all these different websites and products and services. But it's really important to think about that, that so many decisions have gone into the way that communication is presented to you the way that you're given certain choices. And they all have you know, different objectives to engage you or to keep you using this online service or product. And that's really where the power of design shines through. Because designers are making all these important decisions that shape how we experience digital products and services, and they have the power to harm or protect us. So designers can make choices that proactively help us strengthen our privacy, or make choices that, you know, stop us from sharing unnecessary data, or putting us in positions of harm. So yeah, I think designers really have an important role to play and responsibility and how they create those services and help people make these different decisions that you make all the time.

Ben Whitelaw

As a designer yourself, are you. And have you been conscious of some of those responsibilities? In your own work in the past?

Georgina Bourke

Yes, yes. So before joining the ICAO, I was in industry, you know, working with different products and services. I have specialised in data protection in the past. So it's something that I've always thought about, but also been helping other product teams to think about in their own work. It's not something that designers really consider as their responsibility. Often, data protection or privacy is often seen as a kind of legal issue. And not something that's important for designers, but I really see it as completely integrated to the way that you would design a service anyway, you know, thinking about what your users need? And how are you helping them make choices to meet your different objectives? Privacy should just be a part of that. It's part of what we need and how we trust services. And so really, designers are in a perfect position to help people protect their privacy and make sure that's baked in to the design of products and services.

Ben Whitelaw

Can you kind of give an example of where maybe an app or a digital service has made a design choice that is impacted his users negatively?

Georgina Bourke

Yeah. So I think there are there are different ways that design decisions can impact people negatively. Some of those can be intentional. So you can make privacy information overly complex, knowing that people are never going to read it or really challenge anything about it. You can design choices in a way that influence people towards the decision they wouldn't normally make. So I'm sure we've all seen the the you know except or big green button, that you can't really see any any other choices around that. And then you have the more information in tiny greyed out, that's where you actually can turn off, you know, whatever or, or withdraw consent. So there are, there are definitely ways to intentionally lead people towards sharing data or making decisions that they they wouldn't normally make. I think there are unintentional design decisions that lead to harm as well. So, you know, talking about conduct between people on a platform, designers have to be thinking about the tools that they offer people that help them report any harmful behaviour or give feedback. If they've received a recommendation or, you

know, personalised content that is harmful. There needs to be tools in place for them to report that for that to be fixed and removed. And that can often not be thought about or those reporting tools don't work in the best way or don't support the person who's experienced that harm. So I think there it yeah, there are ways, unintentionally, that designers can create services that don't support people or lead to more unnecessary harm.

Ben Whitelaw

Is that a kind of naivety on the part of designers sometimes and people in working in product teams, how is it possible to get into the mindset where you're putting yourself in that scenario where the worst possible thing might happen?

Georgina Bourke

Yeah, really good question. And I think it comes down to this, the fact that designers just don't often think about the fact that it's their responsibility to think about these risks and harms. And a lot of the tools that we've developed, you know, with the Children's code design guidance that are targeted specifically at when are those important moments in design and development, when designers should be considering risks, data protection risks to children, help designers empathise with what would the needs of a child be of this age on my platform or using my service? What are the potential harms that we might need to look out for? And the tools are designed to help product teams think about those different harms and risks that they might not necessarily be thinking about? When you know they're trying to meet the commercial objectives of their product, or whatever it is. So I think it's difficult. But we're creating tools to help product teams think about those harms and risks that they they might not be thinking about.

Ben Whitelaw

And if you're a designer or a product kind of manager listening to this podcast, or you work in a startup concerned with keeping users safe online in various forms, how can those folks kind of engage more deeply with some of the kind of key concepts in in safer design and engage with some of the ICOs work?

Georgina Bourke

Absolutely. So you can check out the children's code design guidance that's on the ICOs website. And this is this is the first iteration of the design guidance. So we're looking to develop it over the next couple of years to include other standards, we focus on transparency initially, because it is such a central part of the children's code and the way that you understand what's happening today to make choices. But it also touches on other other standards in the code as well. And there you can find, you know, visual examples of how certain concepts might be applied in interface design. You can also find practical tools and workshops that adapt common design practices like user journey mapping, or thinking about user needs to thinking about privacy in that context. So there are tools that you can use and, and start adapting your practice to better protect children. And, you know, the whole the whole process of creating the design guidance is collaborative. So as we continue to develop the guidance, we'll be looking to speak to product teams understand how it's impacting them how useful it is to develop new guidance products. So if product teams also would like to become more involved in how

those things are created, we're really open to sort of collaborations there. So please do get in contact as well as using the tools.

Ben Whitelaw

The Children's code is there to give companies guidance on how best to keep young people safe online. It's something Iain Drennan thinks about a lot in his role as executive director of we protect global alliance and network comprised of over 200 members, including representatives from governments, civil society, international organisations, and the private sector, all focused on ending online child sexual abuse.

Iain Drennan

We recognise that this is a global problem that demands a global solution and No single actor has the ability to resolve that by themselves. So we exist to provide that space for collaboration, for knowledge sharing, for amplifying the voices of young people and survivors of abuse, and ultimately for advocacy. So reaching out to senior leaders across our constituencies, and demanding action, what is

Ben Whitelaw

the potential for safety by design to solve for some of the challenges we see around child sexual abuse material?

Iain Drennan

First of all, we're looking at an overall response to child sexual abuse online. So identifying and taking down Child Sexual Abuse material is part of that. But ultimately, that's a response that's reactive, the crime, the harm has already happened. And we're trying to investigate that and remove that. That's incredibly important. However, what we are advocating for is that there is a real pivot towards prevent. So there is more investment and emphasis into stopping the harm before it happens. That is where safety by design comes in. Because when you're at the drawing board stage, you know, when the next tick tock is being designed, or when the next Roblox is being designed. Along with all the technical considerations and engineering considerations, you've got child's safety at the part of it. So our overall vision is a digital world designed to protect children from sexual exploitation and abuse. So that this is a fundamental consideration when shaping and developing this environment. So one of our key members is the Australian office of the Safety Commissioner. And I think they put it very, very well, when they said it's essential to adopt a human centric approach. And I think that's, that's beautifully put, because it rather than just thinking about users as sort of an element in the engineering process, you're thinking about their safety, their rights, and their needs and expectations when you're putting together a process. And they have a great set of guidelines that they've developed in terms of how to actually make that make that happen.

Ben Whitelaw

Can you give an example of a company that is using safety by design, to, as you say, kind of pivot to prevent child sexual abuse material.

Iain Drennan

So tick tock, recently changed their default settings for all under 18 users to make it easier for them to avoid grooming to make it a safer experience for them. So that simple change from the outset in terms of how that service was rolled out. By tweaking that initial default setting, you've immediately got those protections in place from the get go. And we know from looking at nudges, and how default settings can influence those can have major major impacts and users experience. So having that thought process, having child safety, potential risks of abuse in the back of developers minds, when they're bringing products to market or when they're doing updates is incredibly important.

Ben Whitelaw

So just to clarify, what did Tiktok do, for under 18 users in order to prevent grooming.

Iain Drennan

What they did was, if you are a user between 13 and and 16, you have the highest level of privacy settings already enabled, that means users who don't know you don't have access to the videos that you put out. You've got quite a walled garden in place in terms of the default settings. There are slightly different settings for 17 and 18 year olds, but it still offers more protection than the default settings for over 18 users. So I think it's just recognising different needs of those users, and then enabling them to make informed choices. And I think another element that's important to bring out is that the onus shouldn't just be on parents and caregivers or on children themselves. In terms of understanding and making, making choices. Yes, digital literacy is important. It's very important that children know what the challenges are in terms of in terms of coming online in terms of getting the best out of it. But the Not so many part of the response. And I think the metaphor that I, I always think of is, if you're thinking in the counterterrorism space, you know, you go to a sports stadium, and there are barriers that are designed into the fabric of the building. No, they don't look like big ugly bollards anymore. They're part of the fabric of the building. And that just makes a better experience for people going to the stadium, it's less obvious, it's less intrusive. And I think that's what we're looking at here in the digital space where the big ugly bollards, so the equivalent of something goes wrong on a platform, and they've got to retrofit protections on there. The idea of building in default settings, building in protections, whereby it's easier to report right from the get go when you're designing the service means that it's all part of the experience that the child user has, it's not something separate. It's not something added in. It's just part of the holistic experience that you have when you're when you're on that platform.

Ben Whitelaw

technology and safety by design is one of the Alliance's global strategic themes. And yet, much of what we're talking about is an optional thing at the moment for companies. What are you doing to try and change that

Iain Drennan

what we're seeing over the past couple of years is a move towards greater regulation of the online space. So we're seeing that in Australia, they've just implemented online safety legislation, it's coming through in the UK, we're seeing it and Ireland, EU are due to publish their proposals on child sexual abuse legislation, there is a move all around the world to consider what can be done in terms of regulating balancing the online space such that their rights and their responsibilities, we have greater visibility of what the risks are, and that companies feel supported and backed. That obviously has to be

very carefully drafted. So that it's future proof, and it takes into account of upcoming risks of what's going to happen in the metaverse. But I think there will always be a need for voluntary action on top of it. So I think there are no silver bullets in this space, there are no there are no solutions that are going to sort of wave a magic wand and solve Child Sexual Abuse online. That's the very reason we exist because there are different capabilities and different actors required at different stages. However, I think voluntary action complements regulation, because a it provides space to do more than the bare minimum, which we would always ask for. But secondly, with the best will in the world, you are never going to get a global legislative framework that is absolutely coherent across territories. This is a crime, whereby you could have a platform headquartered in one jurisdiction, a victim and another perpetrator and another. And that happens day in day out. And therefore you're going to be dealing with different legislative frameworks. So to bridge that gap. That's where voluntary action comes in. So I think there will always be a place for influencing companies to do more on this.

Ben Whitelaw

On top of companies adopting safety by design principles, governments around the world are also starting to encourage it via legislation aimed at making the digital world safer. In the last few months in the United States, the state of California introduced its own age appropriate design code, and other states and countries are expected to follow suit. It's clear that implementing safety by design principles will be crucial in keeping users safe online, be it gaming, social media, dating, or any of the areas discussed in this series. How it happens, however, remains to be seen. This is the final episode in this series of the safety tech podcast. Thank you to everyone that's listened, and shared insights and feedback. If you've liked what you've heard, do share this series, or perhaps even leave us a review wherever you get your podcasts. If you'd like to hear more about the way safety tech is designing a better future, 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. This has been a for kicks production