“We have to do a better job of putting the burden back on the technology companies to take responsibility”: Julie Inman Grant, Australia's eSafety Commissioner, on Safety by Design and what is needed to effectively regulate the tech industry to create safe online spaces.

Ambika Samarthya-Howard
September 11, 2024

Ambika Samarthya-Howard: Can you tell me a little about yourself and the work you've been doing in Australia around child protection, specifically around digital child protection?

Julie Inman Grant: My name is Julie Inman Grant. I've been Australia's eSafety Commissioner for seven and a half years, and I've served under three different prime ministers from different parties. eSafety was the first online safety regulator in the world, and for our first seven years, there was no other regulator, so there was no playbook. We had to write it as we went along. My legislation, the Online Safety Act, has already been reformed twice and is now in review because we know that technology and threats are always going to outpace policy. Built into our legislation are three-year review periods so we can test the powers and make sure that the schemes are fit for purpose. We started as the Children's eSafety Commissioner, but about three months after I came into the job, they expanded it to cover all Australians.

There's a range of challenges here. One is that this hasn't existed at all in the past. One of the biggest challenges we face is creating an expectation and culture of accountability in the technology industry sector. The Online Safety Act gives me two primary functions. One is as a

national educator and coordinator because we want to prevent these harms from happening in the first place. This is where all the foundational work we've done around research, creating an evidence-base, and co-designing the right kinds of materials that use proper pedagogies that are going to lead to help-seeking and real behavioral change rather than using fear and judgment, which often leads to an amygdala hijack. We put a lot of effort there, and we coordinate efforts across the country. Of course, the other set of functions conveyed to us is around regulation, and we have two sets of regulatory powers.

One is around systems and processes, and the other is around complaints schemes. We are Australia's hotline, just like the National Center for Missing & Exploited Children [NCMEC] is for the United States, the Internet Watch Foundation [IWF] is for the U.K., or the Canadian Centre for Children Protection [C3P] is for Canada. These are backed by 20 years of very strong laws that prevent child sexual abuse material and terrorist content from being hosted in Australia, so almost none of the investigations that we do deal with Australian-hosted content; it's all hosted overseas.

We're part of the INHOPE network. We have four different complaints schemes and a youth-based cyberbullying scheme. If a child is being harassed, intimidated, humiliated, or threatened, and they report it to the social media site and the content doesn't come down, they can report it to us. We serve as that safety net. We have a 90% success rate in terms of getting that content down, and that's mostly by working cooperatively with the platforms.

We have an image-based abuse scheme. It was originally called revenge porn, and I said, we're not going to call it revenge porn. Revenge for what? That's inherently victim-blaming language. Let's call it what it is: image-based abuse. Now, it was very controversial back in 2017 when I said that, but the language and the lexicon have changed. The lexicon matters. We have about a 90% success rate in getting intimate images and videos of Australians removed from the internet. This includes deepfakes as well as sexual extortion, including financial sexual extortion that's largely targeting young men between the ages of 16 and 24. We've seen a tripling of reports there and a 1,300% increase in sexual extortion reports over the past four years.

We also have an online content scheme. This is where we take reports of child sexual abuse and terrorist and violent extremist material. Then we have the newest scheme, an adult cyber abuse scheme. It's a much higher threshold than cyberbullying, but if an Australian adult reports something with the intent to seriously harm and the platform doesn't take it down, we work informally with them based on terms of service. If we can prove there was serious intent to harm, and it's menacing, harassing, and offensive, then we can compel removal notices from the platforms. That gives us a huge evidence base in terms of where the harms are happening and what the new threat trends are, but also where the companies are failing to systematically approach these issues.

We have a set of transparency powers called the Basic Online Safety Expectations. We've now issued 27 notices over the past two years. These harms have covered child sexual abuse material, harmful algorithms including recommender systems, sexual extortion, and terror and violent extremist content. The most recent powers we've used are to test the readiness of social media companies to implement age assurance technologies and assess their knowledge of how many underage users they have on their platforms.

Those are powerful tools. In fact, these tools are being challenged in court today, specifically by X Corp. When we issued the first set of transparency notices to about 13 companies, X failed to comply; we found them non-compliant. We fined them, and they failed to pay the notice. They are challenging these transparency powers, and we're having a judicial review hearing today. We've countersued for civil penalties. But it's interesting because they are claiming that they shouldn't have to respond to a government's transparency reports around child sexual abuse because during the period that we issued the notice, they went from Twitter and incorporated in Nevada to become X. So they say that they're not culpable. We'll see where the court comes down on that. Then we've got codes and standards, there's a whole complex range of things that we are asked to do.

Something was still missing that the legislation did not provide for, and that missing element is understanding how we anticipate technology changes, and how those changes might be either beneficial for our citizens or harmful. This is where the Safety By Design initiative comes in. We have to do a better job of putting the burden back on the technology companies to take responsibility for these issues, to understand the risks and the harms, and to embed safety protections up front rather than retrofitting them after the harm has been done.

This also includes a lot of the work we do around tech trends and challenges. To give you an example, back in 2020, we did one on deep fakes. I could not get the mainstream media to pick up the story because it was a little bit too early, and now you can't read an online article that doesn't mention something about deep fakes. But everything we predicted could and would happen has. Because we thought ahead, in 2021, we were able to make sure that all of our complaints schemes covered deep fakes. Now, we can take action against perpetrators and platforms that are purveying and profiting from misuse of this kind of technology.

Ambika Samarthya-Howard: You come from the private sector, from these technology companies. How were you doing this type of work within the companies? How has your experience having worked in the companies informed what you do in government?

Julie Inman Grant: I've had a really interesting evolution. I never knew I'd end up here. But interestingly, I started out in Washington D.C. working for my hometown congressman from Seattle. So my career has been bookended by roles in government.

I was working on social issues, and one day my congressman looked over my cubicle and said, "We've got this little company in our electorate called Microsoft. Can you work on technology issues, too?" So this was in 1991 and 1992, before there was the internet. Lo and behold, in 1995, I was recruited by Microsoft to become one of their first lobbyists in Washington D.C.

I was on the ground during what I call tech policy ground zero, and shaped the Communications Decency Act, including Section 230. At that time, the players were CompuServe, Prodigy, AOL, AltaVista, Novell, and Netscape. We truly thought that the internet was going to be halted in its tracks if it was overregulated and overtaxed. So that's where Section 230 came from. Even as an industry person at that time, we never thought that that law would remain untouched for 28 years, particularly given how many changes there have been in the industry. Social media was not a thing.

Ambika Samarthya-Howard: What was your main issue serving as a lobbyist for Microsoft? What were the major things that Microsoft was concerned about at that time?

Julie Inman Grant: It's very cyclical. I was working on encryption and the Clipper Chip. I was working on the Communications Decency Act [CDA]. I helped put together the first online safety summit for the White House under the Clinton administration. I was also working on online safety, permanent and normal trade relations for China, and a lot of IP [internet protocol] issues, tax reform, patent reform, and innovation. I was involved in H1B Immigration Reform because we needed top technical talent to build and drive the industries. Back then, we were saying that we were creating jobs, economic growth, and innovation and the government shouldn't be putting brakes on us. For the most part, for the past two decades, governments have really obliged, and this is where we find ourselves now. No brakes were put on the growth or the huge acquisitions made by these companies.

I'd been right in the middle of the Microsoft DOJ Antitrust trial around 1998-1999. I remember seeing that Facebook had acquired Instagram and then they had acquired Oculus and WhatsApp, and I kept thinking that they were basically eating up all of their competitors. Now, not only do we have one behemoth, but Google did the same thing in terms of incorporating. That’s what you start to see when large companies stop innovating. They either imitate or they acquire. So here we find ourselves, and we've got a polarized, hostile environment.

Ambika Samarthya-Howard: How did you move from the lobbyist position to where you are now? Can you tell me what happened next in that trajectory?

Julie Inman Grant: After five years in Washington D.C., I said, "Hey, I'm 32 and I'm single; send me somewhere." So they sent me to Australia, and I started Microsoft's corporate affairs programs for Australia and New Zealand. Then, across the Asia Pacific, I had a safety, security and privacy role. I ended my career after 17 years with Microsoft in Seattle as the head of global privacy and safety. I wrote the company's first trust and safety strategy, and tried to bring all of

the different disparate parts of the company together so we could have a consistent strategy. I tried to bring Safety By Design to Microsoft in 2010, which wasn't the right time. They were becoming more of an enterprise company and were less concerned about consumer harms. What's interesting to see is that 14 years later, they're ratcheting that back and stepping up their online safety programs.

We were very active in the early 2000s on online safety. So 17 years there, two years at Twitter, and then a year at Adobe. My experience at Twitter, although just two years, was very formative. I joined after the Arab Spring, and I truly believed in the strength of social media to be a great leveler to help people speak truth to power. Then, when I was on the inside, I started to see how terrible the online abuse was, and how it disproportionately targeted women and marginalized communities in Australia: those in the Indigenous community, those who identify as LGBTQI+, those from culturally and linguistically diverse backgrounds, those with a disability. I really tried hard to change things on the inside.

I remember being horrified when somebody reported that their nudes were all over Twitter, and Twitter asked them to prove that it was them by uploading their ID. I was like, "No, no, no." This is a person being harmed. Yes, you can ask them to verify that it's them, but this isn't the same as trying to protect a dissident. They didn't have a framework for understanding how to deal with these harms, so one of the things they allowed me to do was create Twitter's first women's safety and empowerment program called Position of Strength. There were a lot of changes while I was there. It was a 140-character microblogging service at the time, but it took a minute and a half to fill out an abuse form. They brought it down to 30 seconds, which was more in line with what the service was, and of course that increased demand for content moderation.

They created things like the muting tool, and they allowed third-party reporting. There started to be some really positive changes. Interestingly, what happened when Jack Dorsey took over again was that he went back to that more hands-off approach and said, "We'll create user empowerment tools." I remember having a conversation with him where I said, "You're putting the burden back on the user." We have a degree of responsibility here to prevent this virality, particularly when network harassment and pile-ons are happening because the person at the end of that attack, of that avalanche of hate, can't be responsible for their own safety when it's being facilitated by the platform. All of these things strengthened my resolve that the technology industry was broken. While they would invest in things like privacy and security, online safety was a very distant priority, even when it came to things like child sexual abuse material.

Our transparency notices have shown that the largest, richest, most technologically developed companies in the world are not doing all that they can to prevent child sexual abuse material. I think Apple's probably the worst offender here. You can see that in the NCMEC numbers. While Meta scans for child sexual abuse material [CSAM] and reported 27 million incidences, Apple reported 267. You can't tell me with billions of handsets and iPads and iPhones out there that there are only 267 incidents of CSAM. The number is 267 because they don't scan. They don't

want to know about it, and they don't provide any reporting mechanisms that are intuitive and in-app for the public to report this to them. They use privacy as a shield to prevent them having to take responsibility for what's happening on their platform, and for what's being hosted and shared on their platform.

Ambika Samarthya-Howard: How did you move into government?

Julie Inman Grant: Interestingly, one of the requirements for the job of the eSafety Commissioner was that you had to have substantial experience in the internet industry or the telecoms industry. That's primarily because the Information Communication Technology [ICT] minister at the time, Malcolm Turnbull, was a barrister, but he was also a technology entrepreneur. He went on to become our prime minister, and he was the first prime minister I served under. He intuitively knew that if you didn't understand the drivers of the technology industry, you couldn't anticipate their talking points before they came in. He understood what they were capable of, but he also understood the limitations, and that the person in this role would not be successful.

Ambika Samarthya-Howard: It seems like a lot of what you did in technology, particularly towards the end, and even as a lobbyist, is very much in line with the same values and principles that drive your work in the government. Do you feel like you have even more impact now that you're working in government as opposed to technology? Where is the nuancing of trying to make these things work?

Julie Inman Grant: I often describe myself as a safety antagonist inside these companies, and I do think I made some progress and helped the executives and the leadership think more critically about the issues and do a bit more.

But one of the reasons I left Microsoft after 17 years– and I had a great 17 years– is that they said, "Safety By Design? We're not sure about personal harms. That's not where our strategic direction is going." They said this even though we had platforms like Xbox and Skype that were both known as vectors for CSAM and online abuse. I realized that I had probably reached my pinnacle there, and I wasn't going to have any more influence.

I would say I had a lot more influence when I was within Twitter, and that was easier to do because I went from a company that had 125,000 people to one that had 3,500. That allowed me to have an outsize impact. I worked with my public policy and trust and safety colleagues, and we got a lot done during those two years. But once I started to feel that commitment to safety was ratcheting back, and I couldn't defend Twitter anymore, I was out. Because you're putting yourself out there. It's your personal brand, too. Then I had a very short stint, less than a year, at Adobe, and that had some interesting challenges, but I actually realized there wasn't a strong culture of corporate social responsibility there.

From day one [as eSafety Commissioner], I worked at the intersection of social justice, online safety, and technology. That social justice piece of the puzzle is important to me. I feel like we've been able to do a tremendous amount with very few resources. Being the world's first regulator of online safety, I've now created, along with Ireland, the U.K., and Fiji, the Global Online Safety Regulators Network. This has grown to about 10 regulators. We've got an MOU with the European Commission and the DG Connect, which is the division that implements the Digital Services Act. We're working closely right now with South Korea on image-based abuse issues, or what they call digital sex crimes and deep fakes.

I've come to the conclusion that no man can be an island, and no regulator can be an island, or on an island, particularly when you're talking about countering the stealth, the wealth, and the power of the global technology industry. I also believe that until the U.S. really starts to hold these companies accountable, the rest of the world is going to have to band together, and we're going to face big challenges. We're already seeing these challenges, as I said before, in terms of some of the more recalcitrant players not wanting to be regulated or really operate according to other countries' social licenses.

Ambika Samarthya-Howard: When you're pushing for change around eSafety from within the government, as you're positioned to do now, do you find it easier? Is it like everyone's already in agreement, and you just have to get the tech companies to do it?

Julie Inman Grant: No, that has not been our experience at all. It depends on the government, and it depends on the leadership that's in place at the time. I think the problem regulators like ours will continue to face is that the expectations for what we can achieve are great; but in terms of the commensurate resourcing, we're all terribly under-resourced when you look at the adversaries we're coming up against. I don't necessarily mean the tech companies because there are a lot of ways that we work together with them, particularly, as I mentioned, with our complaints schemes, but it's also the people out there who are deliberately weaponizing and misusing technology.

When I started, there were 35 people, and we were a $10 million organization. We've managed to grow to about $55 million, with about 215 employees. But think about the scale and the scope of all the websites and the eight different sectors of the technology industry we're coming up against. I've got about 50 investigators who take complaints from the public through our complaints schemes. We're dealing with tens and thousands of reports every year, and we are very effective at having that content taken down. We've got a large technology team. I often say the window to our soul is our website, esafety.gov.au. That's where all of the resources live, that's where you report incidents, that's where you find our programs. We do lots of training for teachers. We do training for professionals.

We're also delivering social media self-defense training for women in the spotlight– politicians, journalists, those in entertainment, anyone in the public eye that's receiving disproportionate

abuse. We're helping them be able to navigate that world to be able to reinforce to them that it's not a fair fight and it's not their fault. They're coming up against organized information operations and gender disinformation. They're coming up against clickbait and recommender systems and opaque algorithms that drive outrage. All of this can be automated and networked at scale and can really have an impact on individuals. We help talk them through what the weapons are, what the attack vectors are, how they build their shields, how they build their psychological armor, and how they assess when a tweet or a post is serious, so they know how to engage and when to engage. We give people those strategies and those tools to better protect themselves.

Ambika Samarthya-Howard: If a government official in say, Guatemala or Thailand, listened to this interview and wanted to put a program together the way you're putting it together, what would be some advice you give them? What would be a few of the insights that you've gained from these seven years that you'd share?

Julie Inman Grant: I think our model of prevention, protection, and proactive and systemic change works. All of that's underpinned by partnerships. We do a lot of capacity and capability building in terms of sharing what works and what doesn't work with other governments. We also have to recognize that not every government, even democratic governments, are going to take the same approaches. The other thing that we've built as part of the Global Online Safety Regulators Network is a group of observers. We're all independent statutory authorities from Democratic countries, but they all have to be arm's length.

There are some countries who consider themselves democracies, but still take a very heavy-handed approach to censorship of political speech and journalistic speech. Our focus and our threshold is around serious harm. We are one of the only countries in the world to do this. Ireland is starting to replicate this and offers complaints schemes. The EU and the U.K., for instance, just use systems and process powers. There has to be a recognition that we're never going to have perfect regulatory symbiosis. What we're trying to achieve is a degree of regulatory coherence. So the first thing that we did in terms of a position statement for the network was one around how online safety is compatible with a range of human rights, and how you want to make sure that human rights are a foundation stone for what you're doing. The second paper that we put out this year was around regulatory coherence.

We are very mindful that governments now are legislating in this area, and I think it was the AI drag race of 2023 that really was that tipping point for governments to start regulating because they were so concerned about the imminent threat of artificial general intelligence [AGI] and the harms that this could cause to individuals, societies at large, and humanity. We're really at the beginning, where there isn't a huge groundswell, but there's now a solid group of regulators who are tackling this space. It can only grow as technology becomes more intertwined in our daily lives, including AI and immersive technologies. We're seeing more polarization in the environment and more threats to individuals and to democracy through misinformation,

disinformation, targeted online abuse, and incitement to violence. We're seeing this spillover into real world harm.

We've had our own experiences, and Vital Voices just did a paper on technology-facilitated gender-based violence that used my case study of taking on Elon Musk around a terrorist video that was going viral. All the other companies complied. They decided not to, and Elon Musk decided to tweet to his 196 million users that I was the Australian censorship commissar and issued a dog whistle for an avalanche of hate. I've been doxxed multiple times. My children have been doxxed. It chronicles the abuse.

You continue to see this playing out in the personal attacks on the U.K. Prime Minister in the context of the U.K. riots, and the demonization and vilification of the Supreme Court Justice in Brazil. The way that online harm manifests against women is very different. The Vital Voices piece, which was done by Nina Jankowicz, actually uses an OpenAI model to parse and examine all the tweets. The second day in, there were about 83,000 tweets. 83% of them were negative, and 10% were directly threatening. E Karen, left-wing Barbie, Captain Tampon, Stasi C word. You can imagine. It's about my appearance, my age, rape threats, death threats, threats to harm my children. Things that don't manifest in the same way against men.

The goal, of course, is to silence me and get me off online spaces and out of public life. I'm far from unique in this position. This happens to female politicians and people from diverse backgrounds all the time. I'm not interested in playing the victim card, but I do think awareness is important, especially awareness around how this manifests differently against women.

One of the key things that we've done is we've worked with a group called Safe Work Australia on what an employer's duty of care is when they require their employees to be online or in the media as part of their job and they receive harassment. What are the rights of the employee to ask their employer for support? Safe Work has designated online harassment as a psychosocial hazard in the workplace, and we see that, particularly with female politicians, their parties aren't supporting them, their parliaments aren't supporting them. They're just told to grow a thicker skin and toughen up.

I met with a number of congressional and parliamentary staffers that are managing the social media accounts, and they're experiencing vicarious trauma by dealing with this stuff and have no support. Things are really broken. We know that one in three professional women in Australia experiences online abuse in the context of their job, and one in four will not take an opportunity or a promotion if it requires them to be in the media spotlight or online. That says to me something about entrenching gender inequality in certain ways as well.

Ambika Samarthya-Howard: I'm curious about the role of shifting cultural norms in everything that you're doing. Are you trying to shift cultural norms from your position in the government? Can you do this work without that shift?

Julie Inman Grant: I don't think you can. I don't think any government can put the burden on a very small online safety regulator to do that. But I think we have had some success in shifting norms. One of the first conversations I remember having in 2017 with a group of young people was about whether or not they had heard of our cyberbullying scheme, and if they report abuse to the platforms, would they report to us? They were like, "Nah, we don't want to be the snitch." That was a strong cultural underpinning. We don't want to be the tattletale, and the companies aren't going to do anything anyway, let alone a government agency. I thought, wow, we really have a crisis if young people aren't going to engage in help-seeking and they're not going to disclose abuse because then they're handling this themselves.

For those first five years, we focused on how we could reach young people. We created a youth advisory council and a social media channel called Scroll with eSafety, with all the content developed by young people so that it's authentic, and we just focused on engaging in help-seeking. In 2021, when we did our Mind the Gap survey, we found that 93% of kids did take action when something went wrong online, including using conversation controls and going to speak with their parents.

The other norm we're trying to change is with parents. Parents are the hardest cohort to reach. The surgeon general just put out an advisory on the stress that parents are feeling, but it's the same. We have free webinars. We go to schools, we've got parent guides in multiple languages. We've got eSafety guides that help them navigate all the apps and games their kids are using and explaining where the parental controls are. We just hammer home every time we get the chance. About 95% of parents say that online safety is their preeminent parenting challenge, but only 10% will go seek out the information. We're saying, "Just start the conversation, start the chat, have it early and often. When you're talking to your kids about school and friends and sports, ask them about what they're doing online. Let them know they can come to you if anything goes wrong."

One of the worst tragedies is around sexual extortion. I just spoke to the coroner in one of our largest states, Victoria, and we went through 17 names of young men who had taken their lives as a result of sexual extortion. One of the suicide notes that the child left their parents was just heartbreaking. It took three hours from the time he was contacted by organized criminals for him to take his life. He just wrote, "I screwed up. I'm sorry. I love you." Our messaging around sexual extortion is, "Disclose, disclose, disclose. This is not the end of the world. Tell someone, come to eSafety, we'll help you through it." You can just imagine any parent would just be like, "Why didn't you come talk to me? I was right down the hall. We could have worked through this together." We're trying to get parents to co-play and co-view.

We just did another study with young people around online gaming. They love online gaming, it helps them with their mental health. They think it makes them problem solve. They want their parents to be interested. We go kick a football with the kids. Let's play a little Fortnite with the

kids too, and make sure they're using technology in open areas of the home so we can see what they're doing.

Almost all of the coerced child sexual abuse material that we see is happening in the bedrooms and bathrooms of homes, where you could hear the parents calling them for dinner in the next room. It's not the same as our parents putting us in front of the TV as the digital babysitter. These are interactive tools. We need to be setting the parental controls, and we're trying to make it easy for parents by providing all this information. I think sometimes, parents put this in the too hard basket. But we cannot leave our children to their own devices, on their devices. We need to be engaged in their online lives the same way we are in their everyday lives. That's the cultural shift we're trying to affect.

Another issue that was very peripheral when I first started was how technology facilitated gender-based violence. We have a crisis of domestic and family violence here, but we know that in more than 99% of such cases, when there is a separation or a child comes into the picture, or a woman gets a new job, or some life event causes a partner or a former partner to feel a sense of desperation, they engage in coercive control and use technology to continue that control and harassment. Often, it's low tech. It might be using an air tag to follow your partner or former partner, or sending harassing texts.

We've been working with the banks across Australia because we started to see that microaggressions were being sent with child support payments. They were weaponizing the online banking transactions. Now, all four banks use our Safety By Design framework and have developed AI tools to assess when financial abuse and technology facilitated abuse is happening on their platforms. They will actually write to the people who are weaponizing their technology and say, "We see what you're doing with our online banking transactions. If you continue, we will sever our relationship with you. You will no longer have a bank account. You will no longer have insurance. You will no longer have a mortgage with us." That is powerful because there's an actual deterrent impact.

Ambika Samarthya-Howard: How did you get that to happen?

Julie Inman Grant: I went in and I found the right people at the banks to tell them what we were seeing. The first bank was Commonwealth Bank, and I was lucky to find a person who cared enough to look into it, and they were shocked at what they found. The thing about the banks is that they're also a heavily regulated industry, and they're doing a lot around preventing scam and fraud, around knowing your customers, and around anti-money laundering. This was just another problem that they were prepared to tackle, and it's provided them a lot of great reputational benefits. That's the other thing I've found. Yes, I have regulation authority and I wield those powers, but it's often the reputational damage and the revenue impacts that really help move the needle in terms of changing the behavior of companies. That, I think, is the biggest incentive.

We've already established that the flawed technology ethos of social media, of moving fast and breaking things, has led to the position we're in now because it's profit at all costs, regardless of the harm. We saw the same thing happening with this AI drag race in April of 2023 last year, where it was about achieving first leader status. We'll fix the problems after we get there. I asked Eric Schmidt at a conference last week, "Are we going to have to continue living with this flawed technology ethos, or is there something that we can do to change the incentives of always being the first mover?" He said it will take strong leadership. When I talk to individuals in these AI companies, they want to do it, but if your competitor isn't compelled to do the same, they consider Safety By Design as causing friction. To slow down a little bit, to get it right, you really have to bring everyone along.

This whole concept of Safety By Design that we introduced in 2018, we did *with* the companies rather than *to* the companies in terms of the principles, the risk assessment tools, and the best practices that were surfacing. That's what we'll continue until we actually see those principles being implemented as a matter of course. For AI, for instance, this means making sure they're selecting the right data, and they're not training this on child sexual abuse material or biased data, but that they're training it in the right ways. That they're releasing this with the right safeguards in place, and that they're engineering out misuse. If that doesn't happen through every step of the design development and the deployment process, and we're not building these tools with the fundamentals being safe, we're going to continue in this endless negative spiral. We need to figure out a way to put the burden back on the companies, like we do with car manufacturers.

It was 55 years ago when Ralph Nader wrote *Unsafe at Any Speed*, and the data around traffic fatalities being reduced with seat belts was implemented. The car manufacturers pushed back at the time, but when you think about getting in your car today, you almost take for granted that there are seat belts, and anti-lock brakes, and airbags. There are all sorts of safety technologies that people seek out because they want to be safe in their cars, and they want their families to be safe. But we still have this concept of technological exceptionalism where we're not holding those companies to account, and that needs to change.

Ambika Samarthya-Howard: What hasn't been working? What are some of the limitations that you've been seeing in the work that you're doing?

Julie Inman Grant: What I call jurisdictional arbitrage and the changes that we're seeing with some of these more recalcitrant companies like the X's and the Telegrams of the world. You can see the governments starting to take action. As I said, today we're in court with X Corp, who's trying to challenge our transparency notices around child sexual abuse based on a technicality in Nevada merger law. They're saying that they went from Twitter to X, and since Twitter no longer exists, they are no longer responsible. I think that will continue to be a problem. There is interesting case law coming out of the United States—the recent TikTok decision, the NetChoice

decision; it's very, very mixed. I was struck by what Matt Stoller wrote about the shape-shifting of technology companies and how when they're *asked* to do something like content moderation, or to remove illegal content, they'll fall back on the First Amendment. But when they *want* to do something, they'll turn to Section 230 to provide them that intermediary immunity.

I think there's a lot that needs to be done for governments to work together to hold these companies to account. I've even suggested something equivalent to a Bretton Woods for digital harms, so that we have agreements and assurance that companies that are domiciled in countries like the United States, where it's not likely that substantive regulation will be published, will respect our laws, particularly when it comes to illegal or seriously harmful content and conduct. And that these companies have a social license to operate. I think that's a challenge.

Then, just last night in Australia, the prime minister announced that there will be a social media ban for children. They haven't determined the age, and whether it'll be 14, 15 or 16. We spent two years working on the age verification roadmap, and we've suggested a technology trial, and that's underway. We're seeking information from the technology companies about what they know about how many underage users they have and what they're doing today. Do they have age assurance technologies and processes in place where we could actually implement and enforce something like this? The big challenge is where governments are motivated to do something. Are they taking a blunt force approach that isn't considering the unintended consequences?

One issue I continue to raise is that we know that LGBTQI kids, as well as kids with a disability and those from culturally diverse backgrounds, feel more themselves online than they do in the real world. They're finding their tribe, and they get a lot of mental health support through the connection and communication that they use online. If we're creating this online forbidden fruit, despite all this work we've done around encouraging kids to disclose to parents, they aren't going to tell their parents when something goes wrong in those deeper recesses of the internet. There are no easy answers here. We need to recognize that there's a lot of nuance, there's a lot of complexity, and that there isn't a silver bullet for any of this stuff.

Ambika Samarthya-Howard: What question do you wish I asked you? Is there anything that I missed?

Julie Inman Grant: I don't think you missed anything, but I would say the government has been motivated and online safety has been bipartisan here. There might be differences in terms of how the different parties want to go about it, but they've created an independent online safety regulator. I feel like I've been at the front of the peloton, pushing up the hill with a lot of shrapnel hitting my face because there weren't other regulators dragging behind this. But also because I'm an outsider; I'm an American who came up through the tech industry, who's doing things in a very different way.

I wouldn't say I've always been embraced by the Australian Public Service because I do come from the outside and I do things differently. The only reason I'm doing things differently is because I want to do what I know will work and be effective. You can't have multiple layers and be bureaucratic and slow. You have to anticipate, you have to be a step ahead, and you have to know how to poke holes in these companies' arguments to be effective because they've effectively been unregulated for 30 years, and they don't want Australia to be the first domino. Once we are, then that groundswell of regulation becomes a global reality.

Ambika Samarthya-Howard: I've learned so much in the past hour. Thank you so much for your time.

Ambika Samarthya-Howard (she/her) is Solutions Journalism Network's Chief Innovation Officer. She strategizes on communications, metrics, impact, product and technology, leveraging platforms for the network and creating cool content. She also leads the Solutions Insights Lab, an initiative of SJN that uses targeted research and analysis to identify and interrogate what's working and what's not in a particular sector or field. She has an MFA from Columbia's film program and has been creating, teaching and writing at the intersection of storytelling and social good for two decades. She has produced content for Current TV, UNICEF, Havas, United Nations Population Fund (UNFPA) and Prism.

** This interview has been edited and condensed.*