Center for Humane Technology | Your Undivided Attention Podcast
Episode 25: Your Nation's Attention for the Price of a Used Car

Tristan Harris: Before we get to the show, we just wanted to welcome all the new listeners who discovered our podcast since watching the new Netflix documentary, The Social Dilemma. The response to the film has been unbelievable. It was briefly number one in India, number one in Canada, number two film in the United States on Netflix and the result for this podcast has been astonishing. As of this morning, we're actually the number one tech show on Apple podcast. We're getting about 100 messages per day from all around the world. In Brazil, Argentina, Sri Lanka, Indonesia. From people saying, "Wow this film has really opened my eyes and it's showing us this thing that I always thought was true but now I can see and understand why it's happening."

Tristan Harris: In a way, there's a little bit of hope in knowing that everyone else is seeing the same problem at the same time. As a friend of mine says, "We're not alone. It's okay, and there's a way out." Speaking of a way out, how are we going to change this massive system? Well, it's going to take change from all sides, just like climate change. You simultaneously need to get people inside of Exxon to be part of the solution and reinvesting all their money into regenerative energy and carbon capture. Just like we want to get Facebook to be as good and healthy and regenerative as possible in the midst of the transition to better and more humane technology.

Tristan Harris: But we're also going to need new platforms and new experiments and social design and how these things can be designed in a way that does not break down truth, that doesn't erode our mental health, that doesn't rely on social feedback loops of approval that end up ruining the psychology of a generation. Governments are not regulating this fast enough. It took us something like six years to get GDPR, the European privacy and data protection regulations, and it's taken more than a year or two to get the California privacy legislation enacted.

Tristan Harris: So given the fact that legislative changes are going to take some time, and these products are still emitting a kind of digital fallout of harms that we know about from the film, one thing that can change fast enough is global culture, what we think and what we do. The people in the film who helped describe these problems are not the only ones who are going to think of the solutions, not by far. This is going to take a mass movement, partnering with the powerful organizations and civil society groups that represent marginalized populations around the world who have been most harmed by these technologies. So what can you do right now?

Tristan Harris: Well, the most valuable thing you can do is to be in discussion with people around you around this problem. We asked you on our previous episode to host a screening of The Social Dilemma with people especially who won't have seen it, who may not have Netflix in their area. It's so much easier now because there's a real global conversation. When heads of state, big celebrities or musicians like Pink or the creators of Game of Thrones or Family Guy are all speaking about it, it's much easier for all of us to bring it up with people who we'd like to see the film.

Tristan Harris: What are the practices and rituals that we could adopt now to live better and in a more humane way, in an inhumane digital environment? We just posted an update to our take control page that has some suggestions for some things that people can do, but I think we need to co-create a better kind of humane doctrine for how we all want to participate in inhumane social platforms whose business models are not

aligned with us. We can have this conversation together. Write us a review on Apple podcasts and leave some of your ideas about what that would look like. What does it mean for you to live in a more humane way in an inhumane internet, and what are the things that you think we should be doing or covering more on this podcast, so that we can help accelerate the solutions?

Tristan Harris: You could ask me or Aza a question or you can tell us what you want to hear on our next episode of Your Undivided Attention. We look forward to hearing from you and we really do read every review. So with that, onto the show.

Zahed Amanullah: We were trying to find out how a targeted message could push back against extremists recruiting.

Tristan Harris: That's Zahed Amanullah, a counter-terrorism expert at the Institute for Strategic Dialogue. He's run a variety of campaigns that are meant to counter the narratives of extremist groups worldwide.

Zahed Amanullah: We tested this out against ISIS propaganda. We tried it with far right groups in the US.

Tristan Harris: Every once in a while, he's startled by his own success. When Facebook gave Zahed's organization $10,000 in ad credits, he started targeting messages in Kenya.

Zahed Amanullah: We weren't expected to reach all the Facebook users in Kenya. It just turned out that after we had spent that money, every single person who had Facebook in Kenya were at least exposed to the ads.

Tristan Harris: To be clear, Zahed didn't quite reach every single user in Kenya. His ads reached 4.4 million people, which was about two thirds of Kenya's Facebook users at the time, but think about it, you can come close to capturing an entire nation's attention for as little as $10,000. That may be a victory for Zahed, but that same victory is now within the reach of nearly any extremist group on a shoestring budget.

Zahed Amanullah: Surely this is happening all over the world where you have this sort of mix of grievance, conflict and extremist groups that know how to use these platforms.

Tristan Harris: They don't need highly produced videos like ISIS, they don't need deep pockets like Russia. With the right message, a fringe organization can reach the majority of a nation's Facebook users for the price of a used car.

Zahed Amanullah: All they had to do was to create accounts, create pages, and look out for those who were engaging with that content. That was all they needed to reach potential recruits. We thought our researchers in Kenya were mainly going to be observing. What they didn't expect was the consistent and repetitive, incessant drawing in through tagging of pictures through personal messages toward invitations to meet offline and some of it was very personal stuff. These people were inviting some of our researchers to family functions offline. They wanted to build that rapport and that familiarity and that camaraderie.

Tristan Harris: This global assault on democracy is not theoretical. Last year, Facebook took down 66 accounts, 83 pages and 11 groups and 12 Instagram accounts related to Russian campaigns in countries on the African continent. Russian networks spent more than $77,000 on Facebook ads in Africa, nearly eight times what Zahed talks about in his interview. As I said in the film, The Social Dilemma, "This is happening at scale by state actors, by people with millions of dollars saying I want to destabilize Kenya. I want to destabilize Cameron. Oh, Angola, that only cost this much."

Tristan Harris: Well, hearing from Zahed, now you'll have the exact price tag. Today on the show, we'll ask Zahed Amanullah head of the counter narrative project at the Institute for Strategic Dialogue, how he fought extremism in Somalia, Pakistan, the UK, and Kenya, among other countries. He'll explain how the very tools that extremists used to broadcast messages of hate can also be used to stop them in their tracks. He'll also explain what tech and government must do to systematically counter the extremists.

Zahed Amanullah: If we don't get in front of this, this phenomenon is going to amplify beyond our reach.

Tristan Harris: I'm Tristan Harris, and this is Your Undivided Attention.

Zahed Amanullah: In 2014, 2015, I think the world was sort of looking at the dramatic propaganda that was coming out of ISIS and Syria. For most people, when they think about extremist recruitment online, they think of those ISIS videos and the production value and all of that. In developing countries, those resources aren't necessarily there but the recruitment is no less effective. In Kenya, the first thing that we did when we went there to analyze what was going on was to map the extremist landscape and part of that was done by getting researchers to explore some of the pages that were being put up on Facebook that appear to be recruiting for Al-Shabaab.

Zahed Amanullah: Al-Shabaab had a long history, going back to the embassy bombings in 2008, when they were influenced by Al-Qaeda and grew in Somalia and were pretty well known by 2014, 2015. Al-Qaeda was still operating. ISIS was relatively new at the time, but Al-Shabaab was the real threat. Al-Shabaab was explicitly recruiting from the Kenyan coastal regions because of the grievances that people in the coastal regions had toward the Kenyan government, and they found them to be ready and willing recruits many of the young people.

Zahed Amanullah: If you engage with those pages, recruiting for Al-Shabaab, if you like those pages, if you share a comment from those pages, that was a window to get at vulnerable people. So our researchers, as soon as they would like a page, they would get inundated by friend requests, by being tagged in photos, being messaged and invited to online meetups and offline meetups. Again, all of this wasn't very explicit. It wasn't, we're Al-Shabaab. Come join us. It was more of a slow grooming process, not unlike that you'd see with gangs or other groups around the world.

Tristan Harris: This mirrors a lot of what past guests on Your Undivided Attention have talked about. Renée DiResta, a Russia disinformation researcher has talked about also how in conspiracy theories ... If you followed on Facebook, one conspiracy theory, let's say, Pizza Gate or something like that, it would say, oh, people like you also tend to

like the anti-vaccine group, the chemtrails group, the flat earth group, et cetera. So one thing I think people underestimate is the power, in terms of the human brain of surround sound. If you join a few of those groups, then you start to get a surround-sound effect where your newsfeed is gradually filled up more and more with more of those memes. I think that the nonlinear impact of surround sound versus just hearing it once from one group is something we tend to underestimate. What do you think about that?

Zahed Amanullah: That's exactly how we were able to be exposed to so many groups over a short period of time. It was a truly immersive experience. We were really worried that there was actual physical risk if we were to go too much further down that road. So we actually had to come up with another way of monitoring content without putting our researchers and our partners in Kenya at risk. We interviewed a lot of NGOs that themselves had a higher risk tolerance than we did and a lot of them were willing to meet people.

Zahed Amanullah: Some of them had worked with former Al-Shabaab extremists. We actually had to ask them to pull back and say, "Okay, we're not going to do this engagement anymore because it's clearly getting out of hand. What we are going to do is focus on the messaging based on your experiences, based on what we've learned from the online mapping, and create something that will resonate with the target groups that are being influenced by these extremist groups."

Zahed Amanullah: We just really just wanted to turn off focus on that. We put a really, really complex picture of the languages that people use, the groups used mostly were Swahili, also Arabic. There was some English content, but that English content appeared to be translated through Google Translate from Arabic source text. Some of that text came from Al-Qaeda. So it was a bit of a haphazard mix of materials, but all with the same purpose.

Zahed Amanullah: Basically, how do we spread this net out using this new tool that is having mass reach and get as many young people as we could to join our movement? So we had this body of data that we put at the lap of Facebook and said, "You have to know that this is happening right under your nose in East Africa? Can we talk to your subject matter experts who are supposed to be looking out for this stuff," and they didn't have any. It's not just, again, not just people who speak the language, but people who understand the nuances that groups use to sort of get noticed, and the trends and so forth.

Tristan Harris: What was that like for you in that moment when you discover all these patterns, and then you bring it to Facebook, and this is in what year?

Zahed Amanullah: 2017, before the elections took place in November. In fact, when we started the project, the first thing we did is to reach out to their teams in East Africa. Google had just set up and Facebook was just putting it together. They didn't even have an office in Nairobi for us to go to. Their subject matter experts were in Palo Alto. We felt it was a huge vulnerability. At the time, we were six months from the presidential election, and we knew that this had a potential to destabilize society in a way that the election 10 years earlier, in 2008, which killed over 1,000 people didn't.

So all of a sudden, this turned from a project where we thought it was just purely academic to something where we thought that lives were really at stake.

Tristan Harris: What's an example of something that you would know if you were on the ground with subject matter expertise versus that you wouldn't know if you were sitting in Menlo Park and having lunch at nice cafeterias and thinking about East Africa?

Zahed Amanullah: Well, the window for us was talking to the dozens of civil society organizations from Kenya, who had first-hand experience getting specific pictures of young people who disappeared overnight and were later found with Al-Shabaab in Somalia, or had escaped from Al-Shabaab, and were captured by the government and disappeared by the government. Piecing together stories about the messaging that they engaged with online, and how that may have led to their disappearance. It was putting that picture together. We were so used to looking for keywords, in English, I might add that we missed that picture. In English, we talk in memes and coded language and dog whistles. Imagine a dog whistle in Swahili. So that's the kind of nuance that we're missing in developing countries around the world, especially where you have languages that aren't universally Spoken like Swahili.

Tristan Harris: I think people tend to underestimate the propaganda power of some of these organizations. Production value might look low to an outside audience. Could you talk a little bit about what people might underestimate about the power of terrorist propaganda?

Zahed Amanullah: When we look at some of the ISIS propaganda that came out in the production value of it, that was the big game changer, that this stuff could be done on a production level that was equivalent to what Hollywood could produce, but you don't need to have high production values to be effective in recruiting. A lot of the stuff that we saw, was actually very crude coming from Al-Shabaab in the region. It didn't mean it was less effective. What's important is that something appears genuine and authentic. Remember, Al-Shabaab is building on grievances that already exist and to do that, they didn't need to be that sophisticated. Their recruitment was and continues to be very effective.

Tristan Harris: And will let us know that it's so effective, even though it has low production value, because I think one of the themes that you're tracking here is just that the cost of doing recruitment is going down over time. The ability to do a highly personalized recruitment is going down the ability to reach exactly who you want, the youth in the specific geographic areas that have the grievances that you want to target. Then the actual literal dollar cost, I think you've said in when we first met that it costs less than $10,000 to reach the entire country of Kenya. If you think about $10,000, might be hard to find in Kenya, but it's not hard to find just about any western country where you want to basically say, hey, I'd like to own the next election, or I'd like to cause grievances or tribalist violence, and how cheap that is.

Zahed Amanullah: It was free, really. It wasn't even low cost. It was ... All they had to do was to create accounts, create pages, and look out for those who were engaging with that content. That was all they needed to reach potential recruits, and surely this is happening all over the world where you have the mix of grievance conflict and extremist groups that know how to use these platforms.

Tristan Harris: So there you are in Kenya, playing with your $10,000 in credits at the same time, you didn't realize there was actually a race to reach every Facebook user from another group which was at the time, Cambridge Analytica. Do you want to talk a little bit about what we now know Cambridge Analytica was doing at the same time that you were doing this work on the ground and with those ad credits?

Zahed Amanullah: Yeah. So when we started, we built up the capacity through a number of workshops in Kenya for organizations to develop impactful campaigns. They are the ones who identified ... They were sort of like public service announcements. We worked with local Kenyan videographers, the subject matter experts on the ground and the storytellers. At the end of the day, we were telling stories, and we wanted these stories to resonate across Kenyan society. There were about 20, 22 campaigns of varying quality, but all authentic and all Kenyan made. So the strategy was to disseminate these videos across all the social media platforms in the run up to the presidential election in 2017 and shortly afterwards. Sort of the zone in which electoral violence was likely to happen.

Tristan Harris: What were some of these PSAs? When you say, PSA, what kind of public service announcement was it?

Zahed Amanullah: Well, the four issues that the NGOs told us that they felt were important to deal with in the atmosphere of an election were combating extremist recruitment from Al-Shabaab, combating the incitement to tribal violence, gender violence, so violence against women and religious bigotry. Kenya is sort of roughly divided between Muslims and Christians, and those tensions often flare up because of ... There's also socio economic divide between the two. So they created campaigns based on those four issues.

Zahed Amanullah: During that same time we have people on the ground in Kenya who were alerting us to other campaign ads that appeared to be supportive of the incumbent President Kenyatta that were very sophisticated, but unattributed, and they were wondering whether they were some of ours because they were ... Some felt quite insightful. We had no idea who was running these ads. We had seen the ads from the official parties who were running for president, but not these unattributed ads.

Zahed Amanullah: It was only a few months later, after the project when there was a sting on Channel 4 News here in the UK, which showed Cambridge Analytica admitting that they were behind the campaign for the incumbent president in Kenya and had delivered ads on Facebook throughout. When that ad surfaced, it would appear to be almost a carbon copy of an ad made for Senator Ted Cruz by a Texas PR firm in terms of the visuals, but it was sort of localized for a Kenyan context.

Tristan Harris: Was it the same Texas PR firm that had been creating those ads for Ted Cruz?

Zahed Amanullah: Apparently it was. I think it was called Harris Interactive or something like that, but they had apparently been the ones to make the ad for the Kenyan election. Now, I don't know that ... I don't think that there was any data stolen from Facebook. The Cambridge Analytica story in the US with regard to stolen data was not applicable in the Kenyan context. It was sort of your garden variety, dirty political campaigning, but nevertheless, they were leveraging the fact that you could reach so many people

so cheap in a Kenyan context and that's what they did, and we were up against that. So our 22 civil society organizations were putting their authentic peace-building adds up against Cambridge Analytica and their insightful ads on behalf of the president.

Tristan Harris: What's amazing to me is that you don't realize that you're competing for that until after you discover these things, and you're actually competing with the opposite messages. Here you are, with 22 civil society groups locally made, Kenyan produced content, and then you're trying to reach all these people. Then suddenly, there's this Texas PR firm out of nowhere working for Cambridge Analytica based in the UK and you realize there's a geopolitical contest that you don't realize you're in.

Tristan Harris: I think that's one of the illusions of the internet is the felt sense of locality. If I go to my Facebook event today and see that, oh, there's an event happening in New York or in San Francisco, and I see 20 friends that are going to go, it feels very local to me. It feels like a very local experience, but little do you know, even you with these 22 groups, that there's this geopolitical contest happening underneath your notes.

Zahed Amanullah: Absolutely. That was really eye-opening for us, because what's the bulwark against that if the social media companies are sort of willing participants, if the government is either unaware or semi complicit? What's the bulwark against that? In Kenya, it was us. That was all we knew it was we had 22 organizations, scrappy organizations sort of getting their stories together and turning them into video content and putting it online. We weren't aware of anything else that was pushing back against that.

Tristan Harris: Also, you discovered this by accident. I often give the metaphor that if Russia tries to fly a plane in the United States, there's a pentagon whose job it is and will successfully shoot down that plane before it gets anywhere near the United States, but if they try to fly an information plane into the United States to sow division, Facebook and Google's algorithms say, "What zip code do you want to target?" So the comparison between a protected zone and a protected stance versus one in which is completely open, that here you are the ones identifying this threat that's coming in, as opposed to even the people in California didn't see this one happening, it sounds like.

Zahed Amanullah: Exactly. That was the big message that came out of this project was that this stuff could happen anywhere. If it wasn't for the sting that where Cambridge Analytica was caught admitting it, no one would have ever known. So people could have gotten away with this over and over again, completely legally and completely without anyone attempting any kind of regulation.

Tristan Harris: I think that's one of the things that was also so interesting to me about your work because on the one hand, you had studied how the tools are used for terrorist recruitment, but then on the other hand, you developed all these counter-narrative positive powers, like using the exact same tools, doing hyper-personalized, anti-recruitment persuasion that says, hey, here's these different way to do it. Then if you map the growth rate of terrorism or hate creation, and then the growth rate of counter-narrative solutions, on the one hand you have these positive examples that I hope we get to about the mechanics of what counter narrative means and how do

you anti recruit people from that, but on the other hand, those things aren't scaling at the same growth rate.

Tristan Harris: So I think one of the things that sitting inside of a technology company right now, you run into this problem, of course, like there's many goods that Facebook does. Then there's these things that we don't feel so good about. People dying, tribalism, election destabilization, et cetera but there's these two balance sheets and it's really hard to compare the goods to the bads. I really want to get people to understand that the growth rate means that there's an urgency to this, that's not being addressed. We're still acting totally reactively to something that is really destabilizing geopolitics.

Zahed Amanullah: That's exactly it. When we did our early research into counter narratives, we were trying to find out how a targeted message could push back against extremists recruiting. We tested this out against ISIS propaganda, and trying to reach the audience that they were trying to reach and see if we could build some sort of resilience or get some sort of positive feedback from the audience, before they became susceptible to that propaganda.

Zahed Amanullah: We tried it with far-right groups in the US, in a pilot study that we did in 2016 with groups like ExitUSA, where we were able to actually get some really, really good results from targeted campaigns. The challenge with counter narratives, and we were very objective about it, the truth is they don't always work unless a whole bunch of things line up. You've got to have a really credible messenger, you've got to have really well-crafted content and you've got to reach the right people and there's an algorithm that we would use to try to identify the target audience.

Zahed Amanullah: We work with organizations like Life After Hate to do that. In their case, it worked really, really well. We did a targeted ad campaign aiming at white nationalist neo-Nazis in the US and out of that campaign, eight people respond to the campaign by approaching the organization and saying that they were in a movement, an extremist group, and they wanted to leave. For us, that was like the gold standard, the most impactful that a counter-narrative campaign could be.

Zahed Amanullah: By the way, in that actual video ad that we targeted ended up winning an Emmy Award. That's how well it was crafted, but most of the time, one of those channels falls short, and you don't have the efficacy that you want. This is where we've sort of evolved from a strategy of countering the narratives of extremist groups to allowing your target groups to build positive narratives about who they are, and about their identity. A lot of this stuff is centered around identity politics, and the ones that are vulnerable to joining extremist groups, they're the ones that are sort of searching for an identity and they just seem to find that in extremist groups.

Tristan Harris: Let's talk about that from them because there's like these different levels of persuasion. You can change someone's environment, you can change their behavior, you can alter their habits, you can alter their relationships, you can alter their beliefs, or you can alter their identity. There's this famous study in behavioral science of, if you get people to rewrite the phrase, I like chocolate, I like chocolate. I like chocolate. I like chocolate. It's one group, and then you test them later on some

... There's some way they do the control effectively, because I know this probably sounds like a trite example.

Tristan Harris: Then they have another group that says, I am a chocolate lover, I am a chocolate lover, I am a chocolate lover. In other words, one is speaking to a belief like I'd like chocolate and reinforcing that. The other is reinforcing an identity, and the one that reinforces identity is far stronger. So in general, when you're doing, a competition, this war for persuasion, you have all these terrorist groups that are trying to win on the war for one identity, which is I am a Jihadist or I am a white nationalist or something like that.

Tristan Harris: Then how do you effectively compete with a positive alternative identity? I heard you also saying that there's a way of letting them create their own identity. So it's not about dictating what that other identity should be, but offering some space to define it on our own terms.

Zahed Amanullah: That's our work with young people generally falls in that category of finding an identity that is positive and constructive and meaningful. We base a lot of our work in the narrative space on how people feel about their place in society and working with other groups and living amongst different cultures. We measure how people feel about that on the back of it, and that to us is a win.

Tristan Harris: Could you talk about some of the programs? I know one was called Average Mohamed. It was helping Somali youth.

Zahed Amanullah: Average Mohamed is a great guy that we've worked with for a number of years, who's ... Again, he created content on his own. He's a Somali American, who saw what was happening with Al-Shabaab, and decided to make sort of tongue-in-cheek content on YouTube about what it means to be a Muslim, what it means to be Somalia in particular. We ... In fact, we tested some of his content in our study in 2016, to see that if he could influence people in his target audience. He was very interested in reaching young Somalis. At the time, there was a number of them who were joining Al-Shabaab from the US.

Zahed Amanullah: So we worked with him on a number of campaigns to see if they could influence that target audience and the one thing that he had was authenticity. He's a very authentic guy who's made ... It's a bit tongue-in-cheek, the content but you remember it, especially if you're from that background and it can work. Some of the other content that we've worked with, we don't disclose publicly, we didn't want people to counter the positive message explicitly.

Tristan Harris: Well, this is the real paradox of all of this is that it's essentially down to what is the persuasion that you trust. So here we had Cambridge Analytica being based in the UK or the United States at the Texas PR firm doing stuff for Kenya. Then we've also got this other guy, Average Mohamed who's doing pro tolerance, respecting religious differences, social cohesion, pro social ... All that good stuff, but then again, based on a remote basis, and the person on the other end of the wire doesn't necessarily know that there is this asymmetric campaign going on.

Tristan Harris: So one of the uncomfortable things about this whole conversation is just, can we get comfortable with forms of persuasion that are pro social, and again, pro social on whose terms, but between a world where you have runaway terrorist recruitment, which results in lethal effects in society and the destabilization of democracies and elections, versus things that are at least trying to create more pro-democratic psychological mindsets of tolerance, diversity, egalitarianism, respect for religious and gender differences.

Tristan Harris: We need to be able to declare that there's some values we're preferring here, and we can't always be transparent about it, as you said, because if you do, it actually gets weird. It backs up the sort of, oh, there's this whole conspiracy theory to get all of us to believe this one thing or another. We have to have a calmer way to recognize that persuasion is ubiquitous, is diffuse, is happening all the time but that doesn't mean that it's all evil, or manipulative. This brings me to my next question, which is about, in your counter narrative work, tell me about the importance of the offline work and how does that balance out with the online? I'm curious if there's a portfolio of online interventions to counter narratives that are pro egalitarian, pro tolerance, and offline work as well. What is the appropriate distribution? How much do you need both?

Zahed Amanullah: Online messaging will always reach more people than offline, but offline content is much more impactful. We actually worked with google.org two years ago to sort of put together a one million pound fund in the UK to get civil society organizations to give us their ideas about what worked. We had like 120 applications for the funding that had some amazing ideas, most of which were offline. There was this one project that was a boxing academy in London that said that they were starting to see a lot of far right sentiment in there ... The people who came there, and they felt they were in a good place to put some messaging or some sort of training to the people who actually came to their boxing academy. I mean, offline stuff.

Zahed Amanullah: It was so successful that we were able to launch, on the back of that, a 10 million euro fund through a Google Impact Challenge that's covering all of Europe to do some of the same work, but also including child safety. Again, most of it dealing with some sort of offline activity.

Tristan Harris: So finding real world offline spaces, whether it's boxing academies or theater performances to somehow embed positive pro tolerance messages and identities?

Zahed Amanullah: And find those local models that can be replicated. The question is, how do you flush that out and how do you make sure that the resources are there.

Tristan Harris: The challenge then, that comes to my mind is what is the coefficient of growth on the offline counters, because we can do some online, but then the offline doesn't scale as easily. So I'm just wondering, if the growth rate of online harm is greater than the growth rate of offline counter positive narrative, how do we deal with that?

Zahed Amanullah: Well, this is where governments should be involved from a city level to a national level. The one thing that you wouldn't be short of is people on the ground who want to participate in those things. It really just does come down to money.

Tristan Harris: You have a strong city's network, is that right? That works with municipal government officials to do this?

Zahed Amanullah: Yeah. One of our strategies to sort of promote this city led approach was the creation of the Strong Cities Network, which we founded in 2015 at the UN General Assembly in New York, and the idea was when it comes to countering hate and countering extremism, usually they're driven from on a national agenda, like from a national government, but cities are the frontline where a lot of the impacts of extremism happen and cities are much more attuned to what's happening. They're much more agile when it comes to responding. It's their communities that are often pitted against each other or threatened by extremism.

Zahed Amanullah: So we wanted to make sure that cities could actually coordinate between themselves and in a way bypass their central governments and really just share solutions with each other. So part of our strategy was to take our research in counter narratives, take our policy research, take our networks of young people, and make them available to this network of cities globally, which includes a number of American cities, but cities in Africa, Asia, all over Europe, and so forth and get them to coordinate between themselves to sort of pick and choose between them what works, and implement them in different environments, and learn from them.

Zahed Amanullah: So we helped facilitate that by making sure that everything that we've learned and everything that other organizations that do similar work is made available to this network of cities so that they can just get on with it and not wait for their central governments to create legislation that may or may not be in tune to what's actually happening on the ground in those cities.

Tristan Harris: This reminds me of what's been very successful in the climate movement is the C40. The top 40 cities collaborating on how to deal with climate change at their local city level and just because cities are such a large powerhouse for both generating emissions, and also responding to things like wildfires, or flooding risks, or hurricanes, things like that. What exactly are you getting them to do? Is it just like a Zoom call with people once a month, or how do you actually make that really effective?

Zahed Amanullah: Well, we started by having global summits, actually. For the first three years, we had global summits where the cities themselves would gather for a number of days to share what they've done. A lot of the cities ... Almost every city that's joined has been the victim of a terrorist attack and has learned something from it. Those things can be shared either in person or they can be shared online. The cities are in constant engagement with each other based on the experiences that they want to learn from each other. So our job is to facilitate that exchange. In some cases, a lot of the cities do programs with each other where cities are twinned, for example, between Europe and the Middle East, or Europe and Africa. We just provide the network for them to engage with each other. So really, what we're trying to do is spur innovation.

Tristan Harris: I want to make sure we cover this paradox that's often cited in the work against hate and extremism. There's this big debate going on about, obviously, in a world of free speech and the technology companies being anchored in a country that is free

speech absolutist, really, above all else, there's a notion of never take down or de-platform anyone. Obviously, there's some shadow banning, like lightweight downregulation of certain folks. YouTube has recently, I think, downranked ... Downregulated some conspiracy theory recommendations by more than 50% as of this month in March, of 2020, but talk a little bit about this.

Tristan Harris: There's this fear that if we take them down, they're just going to go somewhere else and it's going to get even worse, because they're just going to show up on Gab or one of these other alternative social networks or Discord, and it's going to get a lot worse. So what should we do about this? How do we think about this?

Zahed Amanullah: Social media has been the lifeblood of a lot of extremists who are influential. Hope not Hate came out with a study, which showed that the de-platforming of Tommy Robinson significantly reduced his influence, and his ability to reach the audience that he once had. Surely, he can go to Gab, he can go to one of these other platforms, but it's the reach that matters. They don't have nearly the reach that they would have had on a major platform and they know it.

Zahed Amanullah: The ones who are going to step over the line from just trying to influence and then coordinate a terrorist attack, they'll often go to an encrypted channel anyway. That becomes a whole other challenge for counter-terrorist agencies and that's, of course, an intelligence issue. In terms of extremism, and promoting an extremist view, we're catching extremist content on platforms like SoundCloud, and places that you wouldn't normally expect extremists to flock to, but that's what de-platforming does. That de-platforming can continue. Every time we find content like that, we talk to the platforms, and we talk about how they could better monitor the content and remove it.

Tristan Harris: There are platforms as you're saying that do take more aggressive approaches, whether it's the founder of Twitch who's been taking down some of the more aggressive, angry kind of hate troll type players and like you say, there's maybe SoundCloud or Spotify that if you say, "Hey, look, there's this music that's really extreme that you're creating, and amplifying and doing a whole suggested users flow. Here's more white nationalist you can listen to, here's more Al-Qaeda music you listen to." They take them down and that actually doesn't mean that they all go rushing somewhere else in anger and want to take out their knives and guns. There is a value in actually just having values and saying, "We don't serve that here."

Zahed Amanullah: There is huge value in chasing them off the platforms, because they don't all migrate to the same alternative platform. They might-

Tristan Harris: Right, and they can't necessarily reach as many people when they land somewhere else.

Zahed Amanullah: Yeah, they're going to go to any number of platforms. Their ability to coordinate is limited. Just because there's someone on Twitter with 300,000 followers, those followers aren't going to go to Gab. Every time someone is de-platformed, their ability to influence goes down and down and down to an incredible amount. An incredible scale.

Tristan Harris: One of your earlier interviews, you talked about extreme speech mapping and it sounds like some kind of crazy project where you're doing trend mapping on speech and getting early warning systems, and this relates to another effort that I think people don't know that much about, which is the Global Internet Forum to Counter Terrorism, the GIFCT. Do you want to talk a little bit about how do you do this sort of early warning system for hate and the Global Internet Forum to Counter Terrorism?

Zahed Amanullah: Yeah. I'll step back a little bit. The Global Internet Forum to Counter Terrorism was an effort that we tried to influence since its inception, really working closely with all the companies involved. Microsoft, Twitter, Google, and Facebook launched it in a way to create a collaborative body to coordinate resources and strategies to counter extremism and terrorism online. It started with a shared hash database of content, hash database, really meaning digital identifiers of audio and video content that could be shared amongst platforms, so that they could be removed before they were even uploaded.

Zahed Amanullah: That hash database has grown year, after year, after year, which is a great thing but GIFCT is an industry body. It still relies on the cooperation and the input from all of the major social media companies. It needs civil society organizations and governments to all point out where the holes are. GIFCT has been good in the sense that, for example, it can provide resources to smaller platforms like where extremist content has been hosted before, like just pasted or platforms that are run by a handful of people. It's a boon for them, because they can now benefit from some of the resources that the bigger tech platforms can provide.

Tristan Harris: Are there ways that you'd like to see that expanded or, given the growth rate of hate and terrorism online that I think we were both ... I was actually at the Paris Christchurch Call event with Jacinda Ardern from New Zealand, and the thing that came out of the Christchurch Call with Jacinda was, hey, we're going to expand the GIFCT, we're going to expand the hashing that we're doing of terrorist videos, terrorist recruitment, texts, snippets, et cetera. Because there's an area of moral consensus that we agree that this kind of speech that is actually ... Can lead to people dying. We're going to do an early hashing and warning system, and not allow people to run around and create hate everywhere else.

Tristan Harris: Now there's sort of a rising sea lifts all boats kind of phenomenon. Then what's interesting is we're not expanding that set of categories to include the more gray areas of, well, what is lethal speech versus subtly divisive speech. We have to be really careful here, but I think one of the interesting things is as the growth rate of the harms and the risks that are posed goes up, because we're not growing our counter responses in terms of offline, and online conventionally, we've had like, what are we going to do to scale that and is that going to work or what else do we need here? What resources do you need to be more successful?

Zahed Amanullah: That's a really good question. I'll just give you the Christchurch example. I got a call from one of our former employees actually, who joined Facebook describing to me all the ways that the video was being re-uploaded onto their platform to dodge the hash database entries that were being made of the original video. She told me, it was

in the hundreds. Slightly tweaking the videos, slightly changing the aspect ratio, flipping it, adding a graphic here or there in order to get that video up.

Zahed Amanullah: Of course, that coupled onto the fact that it was live streamed, which it was a newer technology that no one has really been able to figure out how to intercept in real time. There was an attempt to do it shortly after, I think, in Poway in California, that failed, but the fact that that was out there was ... Showed us the limitations of just relying on hashes. So there's no harm in the hash database growing but one thing to remember is that hash databases consist of known content.

Zahed Amanullah: Content can be created on an infinite scale in a new way every day. So just because you create new hashes for all the new contents created doesn't prohibit new content with original identifiers from being created on an ongoing basis.

Tristan Harris: I'll say from personal experience, it was about a month ago that someone who studies this showed me a video that was the Christchurch video on Facebook, the upload date was the same day as the Christchurch event, and it still had not been taken down and this is in March, of 2020.

Zahed Amanullah: It happens all the time. We find this stuff all the time and sometimes we're not even sure how it bypassed or escaped the hash database, but this is the new world we're living in. This is why we work so hard on the resilient side of the equation, as much as we do also work on the take-down side, because there is a chance that we'll never be able to close the loop on this, and that the new norm will be this stuff existing in some form online and the response has to be, how do we inoculate society from being adversely affected by it?

Tristan Harris: Well, so this brings me to, I think, our close which is talking about solutions, and it's ... Let's say it's 2030. We've done it. Extremism has grinded to a halt over the last 10 years. We have some version of a social media set of technologies that we still use and empower our lives in hopefully more humane regenerative waves, and we still have a world that relies at least in part on user-generated content. Because essentially, what we're talking about here, when you talk about a society that's more resilient, is the fundamental incommensurability of a model where any of the 3 billion people on planet earth can create, upload, live, brand new content that's never been seen before, which has the potential for harm and amplification with viral scales.

Tristan Harris: That cannot be solved with technology. It can't be. There is no solution to this problem. You need to have a more tolerant and cynical and resilient society that is more aware of this being done, but in a world where you cannot guarantee that each user, each human mind that is being exposed to this uncontrollable Frankenstein of long-tail, user-generated content in hundreds of languages, and not the Swahili but the language no one's ever heard about, and then that's sitting on social media, and there is no AI and there is no hashtag. You cannot guarantee that the person watching that video or looking at that content has even any amount of digital literacy, because all of our trust mechanisms have been mis calibrated.

Tristan Harris: So if we zoom forward 10 years where somehow we did solve this problem, and we did have a more resilient society, and we have done what we needed to do with the tech platforms, in your view, given what you know, what have we done?

Zahed Amanullah: That's some of the most existential of all questions. I ... It's something that I've thought about every year that I've worked at ISD. Are we ready to aspire to that world, or are we doomed to live with it, no matter what happens, no matter all of our efforts? I think the regulation issue is going to be a really, really protracted challenge that requires a whole-of-society approach. Somehow we've all got to get on the same page, across borders, across cultures, across languages, like you said, and I tend to put my faith again, in the innate goodness of people and their ability to build beautiful things in their communities, in their cultures, in their societies and try to nurture that in some way so that the efforts to destabilize that are seen as futile.

Zahed Amanullah: Most people are very, very passive users of this technology. It's just the very few active people that are able to destabilize. To me, it's not a question of extremism disappearing, it's a question of it being impotent.

Tristan Harris: Zahed, thank you so much for coming on the podcast. It's been wonderful to have you.

Zahed Amanullah: Thanks for having me.

Aza Raskin: Your Undivided Attention is produced by the Center for Humane Technology. Our executive producer is Dan Kedmey and our associate producer is Natalie Jones. Noor Al-Samarrai and Mara Kardas-Nelson helped with the fact-checking. Original music and sound designed by Ryan and Hays Holladay, and a special thanks to the whole Center for Humane Technology team for making this podcast possible.

Tristan Harris: A very special thanks to the generous lead supporters of our work at the Center for Humane Technology, including the Omidyar Network, the Gerald Schwartz and Heather Reisman Foundation, the Patrick J. McGovern Foundation, Evolve Foundation, Craig Newmark Philanthropies and Knight Foundation, among many others. Huge thanks from all of us.