

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 20: The Bully's Pulpit

- Fadi Quran: There was this far right candidate, Bolsonaro, who's now the president, but he was a back bencher. The majority of Brazilians said they would never vote for him. And he spoke a lot about basically cutting down big chunks of the Amazon and attacks on indigenous communities and on women.
- Tristan Harris: That's Fadi Quran an activist and campaign director at Avaaz, a platform where 62 million activists gather online to push for change. For years these activists have been warning Fadi that disinformation is hindering their ability to organize. Whether their cause is climate change or human rights or vaccinations they can't seem to get people to agree on basic facts. So Fadi and his team began studying how disinformation poisons the atmosphere for activism. That's why they had a front row seat to the Brazilian elections in 2018.
- Fadi Quran: Our members in Brazil were some of the first who began saying this thing is becoming more serious. And although six months ahead of the elections in Brazil about 66% of Brazilians said they would never vote for this guy, what we began seeing was that more and more people were beginning to support him as the social media environment became more and more toxic in Brazil.
- Tristan Harris: The decisions made in conferences rooms in Silicon Valley reverberate throughout the world. We've talked about this before on the podcast but our guest today can really show us those effects and tell us about the people who experience them firsthand.
- Fadi Quran: They're just sitting there being pummeled and pummeled and pummeled by this horrendous content that's making people literally go out and look for them in the streets and beat them up and bully them.
- Tristan Harris: In the last few weeks we've watched as anger, frustration, indignation and so many other emotions that have been building for centuries in the United States rose to the surface of our national consciousness. At Your Undivided Attention and the Center for Humane Technology it's felt like a time to listen and to reflect. We actually held back many episodes that didn't meet the present moment. But we think now is the time to put this interview out there because social media has played a major role in the Black Lives Matter Movement and the coronavirus pandemic.
- Tristan Harris: Of course, it hasn't been all bad. It has exposed incidents and issues that most of us would otherwise have not seen. It's enabled people to come together into the streets. But the harms have also been enormous. It's also enabled mass hate and the people with the power to reduce the harm aren't close enough to feel the pain. To feel the urgency to fix it. We hope that this interview gives you a better sense of just how global the reach of these harms is and what we can do to repair the damage.
- Tristan Harris: But we've seen signs of progress on that in the last few weeks too. Some Facebook employees have staged a virtual walkout in response to Mark Zuckerberg's decision not to enforce content moderation policy on posts that incite violence. A few have even quit. More than 140 scientists from the Chan Zuckerberg Initiative said in an open letter to Mark Zuckerberg that Facebook's practices are directly antithetical to CZI's goal of building a more inclusive, just and healthy future for everyone.

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 20: The Bully's Pulpit

Tristan Harris: We all deserve tools that not only bring us together but also keep us safe and bring out the best in us as humans. We are lending our platform to elevate more voices as we listen and learn. Fadi's is just one of them.

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

Tristan Harris: So Fadi, thank you so much for coming on the podcast. Why don't you give people a little bit of background on Avaaz and some of the work you're doing and then we can lead people into a deeper conversation.

Fadi Quran: So, the best way we describe Avaaz is as a movement of people all over the world coming together based on the belief that we have more in common than that which divides us. And most of our campaigns focus on issues such as human rights in the Middle East, climate change, pushing for democracy.

Fadi Quran: But recently what we realized was that the disinformation environment and the larger social media environment was creating an ecosystem that made it hard to achieve the goals that would make the world a better place. And our members, particularly in 2016, 2017, began urging us to work on this issue. And the people who we were in touch with, whether they were communities in Assam in India or women's rights movements in Saudi Arabia or average U.S. citizens who were facing different health crises, were telling us that this problem was becoming more and more, not just serious, but becoming more of an existential threat to even the small campaigns that people were organizing on the ground.

Tristan Harris: And so how did this first get on your radar from some of the communities who are most affected.

Fadi Quran: So, we do a lot of work in Brazil. About 10 million of our members are based there and we campaign a lot on issues such as defending the Amazon and the indigenous communities that are there. And as the Brazilian elections kind of began moving forward we built a team to begin monitoring this and one of the things that we found is that about 89% of those who voted for Bolsonaro eventually in the elections had believed one of the top 10 crazy stories that were put out against his opponent. And this just shocked us.

Fadi Quran: And then we began looking in other places. So one of the stories and issues that we campaigned on when the genocide was happening in Myanmar. Our community mobilized millions of dollars to help those who were running away from the genocide. To cross into other countries and to get away from it. And as we were working with these victims on the ground in Myanmar a lot of the stories that were coming out of them also were indicating to us that social media, particularly Facebook in this case, was inciting violence against them. And the stories that we began hearing were stories of, for example, a woman refugee whose husband was slaughtered in front of her and who was raped, because, we're not sure but it's likely the government had spread pictures of her husband and their village. Saying that they were terrorists and they were planning attacks on the nearby communities.

Fadi Quran: And so, as we began to hear these stories and then as we also began to see things such as Brexit and what happened there, Jo Cox, may she rest in peace, the member of

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 20: The Bully's Pulpit

parliament who was murdered there was a close friend of mine and many of the people on the Avaaz team. And the person who killed her was clear that social media at least played the role in further radicalizing him and his beliefs.

Fadi Quran: And so all of these different stories began to come together to indicate that something's really wrong is happening. As we dug deeper and as we polled our members we also began to realize it's not just about these communities. It's the regular kid next door who may be being bullied at school, who suddenly finds himself sucked into this information environment that's just full of false information. And is either becoming socially isolated or becoming more and more violent due to what they're reading.

Fadi Quran: And the existential threat comes in at two places. The first step is the kind of direct dangers that we see happening to individuals, whether it's what you speak a lot about, the different kind of rabbit holes that people can fall into in terms of radicalization. But also, it makes it difficult for us as humanity to connect and find deliberative, nuanced, intelligent solutions to challenges such as climate change. Having access to facts, having access to a community that can engage on these issues with complexity and with hope. And that means that we'll be in very, very deep trouble. That could, and I'm not exaggerating here, that could and I believe will threaten the human race.

Tristan Harris: Yeah, I agree with everything you've just shared. I was just in Washington D.C. meeting with and briefing several senators and one of the things that they were talking about is that they go home to these town hall meetings and people are believing crazier and crazier things. You know, oftentimes we've said that this is the problem beneath all other problems, because if we can't agree on what's true everything is gone. And so right there with you that I think people need to hear and understand step-by-step why it's an existential threat.

Tristan Harris: I think the other example you mentioned of Brazil is so fascinating and alarming to me and I think it would be great if you could share more about it, because it's only two steps from social media. Like you said, before social media people were saying they weren't going to vote for Bolsonaro and then suddenly, you say 85 or something percent of people who voted for him believed one of the fake news stories that had been spread about him. When he gets elected and then he makes decisions to start burning a huge chunk of the rainforest, which is the lungs of the earth, it's only two steps from social media electing a far right leader, which then leads to changing the entire future of humanity, because he's hollowing out the lungs of the earth.

Tristan Harris: So, that's an existential sort of decision-making cascade right there. Could you say anything more about what you saw happening in Brazil?

Fadi Quran: Yeah, definitely. We built this kind of small team of elves and we just began collecting what was spreading on social media and flagging the different pages and actors that were spreading it. And we realized very quickly that this was coordinated activity by these malicious actors, by Bolsonaro and of his allies, and more importantly, that their abuse of the algorithm was in itself then being amplified by how the algorithm is already built. Whether it was Facebook or YouTube, which also played a big role.

Fadi Quran: And essentially, the narrative slowly began to shift in the country. You know one thing they often teach in political science classes is, in any elections people at max can

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 20: The Bully's Pulpit

remember three things about the candidate that's running. So if you can just basically stick one negative idea into a voter's head about a certain candidate, you can shift their voting behavior. And if you can then also shift the narrative of the elections, what the elections are about, then you can win that election.

Fadi Quran: And first of all, we began flagging some of these pages to Facebook and other social media platforms. The pages that they took down that we flagged had about 12.5 million views and for an election in Brazil that's significant. But then we wanted to see, does the social media environment actually influence elections?

Fadi Quran: So we ran this poll. We chose the most viral stories that had been flagged by fact-checkers in the country, and then we just asked people, "Of these stories, did you see any of them and did you believe any of them?" And it was a shock to us that the voters in Brazil that voted for Bolsonaro, of that subset, 89% of them had believed at least one of those fake news stories. We also found that WhatsApp was being used to then spam people with this disinformation.

Fadi Quran: And it's important to understand how these different platforms, Facebook, WhatsApp, Twitter, YouTube, interact with one another. Where these actors are so sophisticated that they can use Twitter to test very quickly what memes catch on, what disinformation catches on. And then they can move that to other platforms such as Facebook. Create groups and saturate them with the most effective memes and have them share them across society. And then they can use kind of WhatsApp when those memes are let's say, caught by fact-checkers or downgraded, to further spread them into society. And you create this vicious loop that in the end will allow for dictatorial leaders, such as Bolsonaro, people who have these far right beliefs, to win elections.

Fadi Quran: And I think the next important step here that we need to discuss is how these platforms become the best weapons for bullies. If any bully who would want to design a platform where they can bully the most marginal, the weak, the average person, these platforms are just the best tool to do that. Because they can use tactics such as censorship through noise, such as threats and such as outrageous content. And to just keep slamming people with it until basically, you have this kind of triggered reaction from people where they either become despondent, depressed, they don't feel they can impact politics, they don't believe in institutions, they believe everybody's corrupt. Or you move them towards the direction of becoming more and more radical, far right and extreme.

Fadi Quran: And if you do that to a society and you end up getting a leader of basically the most important country in Latin America, Brazil, to then move on policies such as taking down the rain forest, the Amazon or attacking women's rights or attacking indigenous communities, then the society begins to slowly collapse. Unless we, of course, act together to break this kind of hold and rebuild things in a more hopeful way.

Tristan Harris: So why are some of these places more vulnerable to this? I think people coming from a Western audience where we have reasonably functional institutions, many different media sources. Paint a picture of what might be different in some of these places with regard to the strength of institutions or how much their media is really impacted by social media.

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 20: The Bully's Pulpit

Fadi Quran: Yeah. So, an important caveat to my answer here, and I want to keep re-emphasizing this point for our listeners is, we're all vulnerable. Because there's this assumption that disinformation, toxic social media, is more harmful in some places than others or a lot of the listeners may be like, "Yeah, but I would never fall for that. I would never fall for that fake news." But the truth is all of our societies are vulnerable. And our responses, the responses in some places, like on health issues for example, the response in let's say the U.S. may be better than the response in Brazil, because U.S. health institutions may have more resources, more access, more doctors and so forth.

Fadi Quran: But in terms of vulnerability to it, I think what we're seeing across the globe firstly, is that a lot of the news of course environment where people get their news from, especially in younger generations, is coming through social media, because that's where people are spending a lot of their times. But in Brazil specifically, Brazil actually has a relatively healthy media environment. It's not as bad as a place like Myanmar where you would have government-controlled media.

Fadi Quran: But in Brazil, what you do have is cases of corruption that were in the news consistently, where people didn't trust their politicians. That trust was broken already because of the mass scale of corruption that was being reported on. So people began to just look and want to engage the news from other sources, whether that be YouTube or Facebook. The average person in your community and what they post.

Fadi Quran: And that becomes dangerous when people make the wrong assumptions that our politicians and our media is fully corrupt. They kind of extrapolate from cases to define the whole environment that they're in as fully corrupt. And then they make the wrong assumption here is that they make the assumption that, "But what we're seeing on Facebook, what we're seeing on YouTube is the truth." And what will happen then there is, and Tristan you speak a lot about this is, when people begin to use these platforms more and more, the kind of recommendation algorithm, the news feeds, begin to move people towards, because of the addictive model, towards more and more outrageous and radical models. And so people begin to be sucked into that environment.

Fadi Quran: The other thing you have is, in Brazil, Brazil's a big country. It has some urban areas but it also has a significant amount of its population in rural areas. And in those rural areas the kind of local media wasn't as strong, wasn't as big and most importantly, wasn't resourced enough to move quickly to kind of how to use social media to spread their messages. And that created the dynamic there where, you can imagine... I'll give you a story here.

Fadi Quran: There's a woman, whose name is Fabiana de Jesus and there was a fake news story spreading about a woman in one of the villages in Brazil who was kidnapping children and hurting them. And there was a sketch of this woman. And the story went pretty viral in this community and one of the guys saw this mother of three walking down the street and he thought that she looked just like the picture of the woman that was spreading on social media. And she was murdered. She was beaten to death in the street. And because you don't have, in those communities, an effective police force that could quickly engage and shut down this type of fake news media, because you have this belief that the official media is wrong and what we see on social media is right, it led to the death of this mother.

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 20: The Bully's Pulpit

Fadi Quran: What's happening now, even in the U.S., as we look at the U.S. political environment, it is pretty similar in terms of how people are engaging with some local fake websites or fake local news sources. Believing it more than they believe you know, there's this attack on the official media. Fake news, all that type of rhetoric and they're beginning to believe these fringe outlets that are moving people more and more to the extreme.

Tristan Harris: Yeah, it's interesting because we talk often about the power of social media to generate Me Too Movement or positive activist campaigns, Black Lives Matter, even Avaaz is obviously an organization that uses social media to try to generate positive change. And I think an issue that comes up often here is the temptation to say, "Well, there's all these goods and then yet, there's all these bads." And then we kind of throw our hands up and say, "Well I guess we just don't know. It seems like it has goods and bads, let's just leave it alone."

Tristan Harris: And I think the thing that brings your perspective and mine together is that if you look at the balance sheet of harm and the harm actually having a secondary effect of debasing the entire trust in the society. That can send you into a kind of dark ages where people mostly trust the people around them for what's true because they don't trust any of the media sources. And they trust social media, what people are sending them. That balance sheet of a digital dark age, people not knowing what's true at all or being apathetic or being unable to take powerful positive actions, as opposed to being passively outraged yelling at your screen, that is an existential issue.

Fadi Quran: Yeah. It becomes an existential threat because you begin having measles outbreaks. You begin having other diseases spread, and we saw the same thing in the U.S., and that puts everyone's children at risk. I also want to mention a point connected to an inability to have sophisticated conversations.

Fadi Quran: So Facebook and Twitter and others, they say that if people report misinformation or hateful content then they have teams that look at that. But there's a whole other argument about how effective these teams have been. But let's just say that these teams work perfectly. The truth is, you have a lot of communities who don't have access to social media, who don't have access to technology. The best example is the indigenous community in Brazil or the Muslim community in Assam in India who are much poorer, who have less access to these technologies.

Fadi Quran: So you will have this cycle, this bullying cycle essentially, where, even for people who are not online, they're being attacked and targeted online, on social media, by these bad actors. And they can't report it. They're not online. They can't respond to it and spread their own stories. They can't even have a conversation about it, they're just sitting there being pummeled and pummeled and pummeled by this horrendous content that's making people literally go out and look for them in the streets. And beat them up and bully them because they're being called rapists. And they just have no response at all.

Fadi Quran: So there's this assumption that social media, and this is an argument that we hear a lot, that social media opens the conversation for everybody. Everyone has freedom of expression, as if it's this perfectly fair environment. But it's actually skewed. Not only is the algorithm skewed but the privilege of being on social media is skewed. And skewed to benefit bullies and that kind of slope is just getting more and more intense.

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 20: The Bully's Pulpit

- Fadi Quran: And you know I started a lot of my work in activism and you mentioned also the work we do at Avaaz. When I started organizing online, in 2011, around the Arab Spring for nonviolent protest, I remember being arrested by the intelligence here. And they took me in and they made me open my laptop and they said, "You're going to delete this Facebook event now." They had no clue what Facebook was but they were like, "You need to delete it." And I realized they didn't know what they were talking about at the time. So I said, "You can't delete an event off Facebook. That's impossible, I'm sorry." And they were like, "Well you're going to have to put a post that says that this event is canceled."
- Fadi Quran: So I looked at them and then I just wrote a post saying, "Sorry, the intelligence are at my house and they're demanding that I cancel this event." Knowing that it would actually lead more people to come to the street. And then they were happy with that and they released me. But they had no understanding at the time. And at that time it did benefit organizing. Online social media benefited that.
- Fadi Quran: But today, we're in a whole different environment where governments, they're the ones that have the power and privilege and control and access and resources to really define the debate online through fake accounts, through a million and one means. Political advertising now, which Facebook decides not to fact check. So the whole environment is just skewed more and more towards authoritarian powers. And it's not the social media that was promised. It's not the social media that we hear about on these about pages of these platforms. It's not even serving the mission of the social media founders anymore, it's serving the exact opposite mission, which is allowing powerful bullying actors to shut down the debate using many malicious means.
- Tristan Harris: Yeah. It's so hard to hear about these things. I'd love, Fadi, for you to talk about a meeting that you had in May, 2019 where you brought an incredible set of people who have been affected by these problems from around the world. And flew them in from Myanmar and Finland, New Jersey and Sandy Hook to actually talk to the tech companies. Could you talk about how, when you present these issues to tech companies, they've responded in those meetings?
- Fadi Quran: Yeah, definitely. Yeah. So, to take a step back here, I studied at Stanford so a lot of, as did you, and a lot of the people now working at these companies are people that are my friends. People that I've been close with. And one of the things that I and our team realized is there's a disconnect between the kind of work environment and what the executives and employees at these companies believe that they're doing, believe that their platforms are doing. There's kind of a narrative about that and between what's actually happening to people as a result of these platforms and their algorithms.
- Fadi Quran: And the idea was, they're still good people. The people I studied with, the majority of people that work on these companies, are not bad people. And our hope was that human connection, the thing we want to build, the thing I think you and I and many of those listening to this podcast believe, is that human connection can actually heal the world if done well.
- Fadi Quran: So we spoke with a number of key victims of disinformation for example, Lenny Pozner who's a father of one of the Sandy Hook children who was murdered. And who until today, eight years later, is in hiding because of the disinformation that was spread against

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 20: The Bully's Pulpit

him. Number one, just think about that. This wonderful man and his family cannot go out in public because so much filth and malicious lies were spread about them on social media that they're under attack and harassed by gun lobby supporters who basically believe that he's an actor and the whole Sandy Hook thing was an act. And he has a powerful story.

Fadi Quran: We spoke with a leader from Burma from the Rohingya community named Tun Khin and Jessikka Aro who is a Finnish journalist who uncovered Russian disinformation and became the victim of troll attacks and threats on the street. Ethan Lindenberger who is a wonderful, wonderful young man, high school student whose mother believes in anti-vaccination theories. And who has tried to convince her to vaccinate his brother and their family, but she has been sucked into that toxic environment.

Fadi Quran: And we brought this group together and we said we're going to go and meet with executives at these companies, and product designers and engineers. And when we walked into the room and they just shared their stories at Twitter, Facebook, Google, people in the room were in tears. You could see them shaken to the core by what they were hearing. And they couldn't argue that these were just fringe issues because they saw the pain on the faces of those who were in the room and they heard their stories.

Fadi Quran: And we saw that it had an impact. YouTube, for example, after our visit decided to take more stringent action against Sandy Hook conspiracy theories on their platform. Twitter removed one of the key war criminals who designed the genocide against the Rohingya. And we didn't get to meet, I think this is key, the CEOs and that's why we're planning another survivor's visit to Silicon Valley within the next two months, and to D.C., focused on getting these stories to the right people, because we believe that they can have impact.

Fadi Quran: And for everyone listening to this podcast, I think one of the ways to fix this problem is actually to get those who are working on these platforms, those who are designing their interfaces, to just see the stories and consequences of what's happening on the ground. To normal people just like them, because of these monster platforms that have been created today.

Fadi Quran: But I also want to say another thing that's key in our discussion with many of the people in the room. And these decisions for the most part, are not being made by the normal employees at these companies. For those living in California, for example, the threat of climate change is clear. Last year millions of families couldn't take their kids to school because the air pollution was so bad due to the fires in California. So people not only understand this threat, they feel it, they know the danger and they want to do more. They want to make sure that the companies they work at do more. They don't want to be a part of the problem, they want to be a part of the solution.

Fadi Quran: But you have executives at these companies that I think deserve to be called out, because they're often, based on again, what we've heard from meetings inside these companies and what we've heard from speaking with different leaders around the world who've met with them. And just from what we're seeing on the ground, who know that they can do more. Who have solutions on the table such as correcting the records, such as detoxing their algorithms, such as creating more humane algorithms. But they're not implementing them, I think for two reasons. Oftentimes it's out of fear of the political

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 20: The Bully's Pulpit

repercussions, of the exact type of leaders that came to power because of the toxicity of these algorithms. Whether we're speaking about Trump or Moody or Bolsonaro or others, and B.S. in Saudi Arabia, others around the world who are toxic leaders that came to power through social media. At least in a big part and who these platforms don't want to challenge, largely because they may be afraid of anti-trust action or because they may have certain people within the executive suite at the company who share certain interests with these leaders.

**Fadi Quran:** And it's time, I think, for again, employees at these companies and for these CEOs to come and hear the real stories. To engage with the people who are the victims of disinformation and hate on their platforms and to begin acting as fast as possible in the face of this existential threat.

**Tristan Harris:** You're making me think of a phrase that someone brought up in one of these meetings that I've attended, which is, that the people closest to the pain should be closest to the power. And what you are doing when you are bringing these people into the room, and I'm curious to hear more about it, is you're closing a loop that is unclosed. There's an open loop right now between, "I make a decision," and the scale of a designed decision or a tweak in a parameter of which newsfeed stories come to the top or the bottom. Or which people I recommend for Twitter followers to suggest a new user for them to follow.

**Tristan Harris:** Each of those tiny decisions actually impact millions, if not billions of people and one of the problems is that the people who are closest to the pain of those negative impacts, there're so many of them. And they live miles and miles away and they speak, often, different languages than the people who made the decision in the first place. You know we often say there's at least 22 languages in India, how many engineers at a Facebook or YouTube or Twitter do we think speak the 22 different languages of India?

**Tristan Harris:** So what happened? I heard that people were touched and moved in the rooms of tech companies that you brought them into, but you know you said some changes occurred. But oftentimes it takes a lot longer for that change to occur, doesn't it? It's been years in some cases that these things were reported and then didn't get addressed at all until you bring people into the room like that. If you can talk a little bit about what specifically you were asking for when you met them?

**Fadi Quran:** Yeah, definitely. So, I agree with you completely on closing the loop and that the people most impacted should be the people closest to the decision-making. And what I would say is what happened in those meetings is there was movement. I think we mobilized a good number of those engaged on these topics to act more forcefully and pursue more forceful solutions. I unfortunately can't go into detail but we do have promises from all these platforms that they're testing and they're going to roll out much stronger solutions to these issues. And particularly in this case, we hope that they do it as soon as possible because of the U.S. 2020 elections coming up. I wish I could say more but we did get promises.

**Fadi Quran:** But you're right that so far not enough has been done. And our key asks in those meetings revolved around two main goals. The first is what we call "Correcting the Record" and the idea here is that when a person is targeted with disinformation, with false or misleading content, and that's flagged and fact-checked by independent fact-

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 20: The Bully's Pulpit

checkers, the platform should alert everyone who was targeted with that misleading content. To say, "Sorry but you viewed harmful or misleading content or disinformation," and then provide them with sophisticated links, because it's important here not just that you provide the one-line correction. It's important here that you provide sophisticated, well-designed corrections that will engage people and be memorable.

Fadi Quran: And we've done testing on this internally at Avaaz and we also have two academics that we've been working with at Ohio State University and at George Washington University. And we've shown that by showing people well-designed corrections you can decrease the amount of people who believe in disinformation by up to 50%. Sometimes up to 70.

Fadi Quran: The second ask is what we call "detoxifying the algorithm" and this is, I think, very closely connected with a lot of the work that the Center for Humane Technology does but it's about how do you redesign algorithms to be more humane? To put human rights standards first. And mainly to stop recommending disinformation and hateful content to users. And so, for example with YouTube, one of the key things we think YouTube could implement today is, if there's a channel on YouTube that is spreading, let's say climate denial misinformation and this channel is consistently spreading this misinformation. Our research found that YouTube is actually recommending it. It's sending millions of people to these types of videos that then make it hard to reach climate policy.

Fadi Quran: But what YouTube can do, at least as a first step to detox its algorithm is, let's say there's a three-strike rule. So a channel is found three times to have shared misinformation content that is harmful to society, with the purpose to mislead people. Then YouTube can say, "We're going to stop recommending videos from this channel to our users." The videos can stay up there. As Avaaz we have fought in countries like Iran for freedom of expression and stuff, but they shouldn't be recommended to users. Users shouldn't be pushed down that rabbit hole. So these were the kind of two big things that we were asking for in those meetings.

Tristan Harris: And that mirrors the approach of, "Freedom of speech is not the same thing as freedom of reach." That you're granted a right to post content on the internet but there's nothing that says that we each have human rights to reach millions of people just because we open our mouth. And so, yeah I hear you there.

Tristan Harris: And also, it's interesting, the researcher Briony Swire-Thompson's research at Northeastern, she found that... really interesting looking at how human memory works. That when you issue a correction you can't just say, "This thing that you read, that's not true." Your mind doesn't work like a whiteboard eraser where you just erase the other thing. You actually have to provide the factual alternative. The corrections work better when you repeat the retractions multiple times, because our automatic memory is based on salience and repetition.

Tristan Harris: And so, one sort of other additional thing would be, "How do we repeat that correction more often than the original thing got airtime?" There is actually a suggestion from one of our listeners to this podcast that every time there is hate speech that's taken down, we notify the person who posted that hate speech that we're going to post three times

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 20: The Bully's Pulpit

as much positive information about the thing that they were hateful about. And that not only changes and corrects the automatic memory in terms of what gets repeated more often, but it also creates deterrence because I know that if I communicate in the future with hate speech, I'm actually just going to get drowned out by the opposite.

Tristan Harris: And another thing that she found in her research is, if you tailor the correction to the audience. So if you have more personalized things, because if, for example, you present corrective information in a way that's threatening to the person's identity, then actually it will backfire, obviously. And so finding an unthreatening, more strategic way to personalize the correction that's another big one.

Tristan Harris: But you know I find this so interesting because getting this right has less to do with getting the right algorithm or the right tech or the right machine learning or the right data. It's about the human mind. It's about sociology. It's about how we work and I think you and I both share an optimism in a way, even though people are hearing lots of criticism, that we believe in the power of human connection and basic decency and civility and goodness. But you have to design systems that make sure that that's what's brought out in us as opposed to unregulated, unchecked systems that mainly have a lot of bad actors and authoritarian actors. And the most funded nefarious actors to seriously out-compete everyone else.

Fadi Quran: Yeah, 100%. And we see this every day. We don't see enough. People are good. Everyone to a large extent, loves their children. Everyone wants to make sure that their families and friends have a healthy lifestyle. No-one wants to see their neighbors drowning in solution or having their house burnt down. And what humanity has created through this power of connection in the last decades and centuries, is something really beautiful.

Fadi Quran: And if you look around you now at the team that's working with you, if you look at Silicon Valley and what has been created in terms of technology for the world, there's so much beauty in it. And it does show you that we feel agency towards wise interactions. We feel agency towards serving each other. And I think that's what gives me hope and that's also what just terrifies me about the environment that social media has created for our minds and behavior.

Fadi Quran: I was reading a book by Sapolsky, I believe. A professor at Stanford. And he explains in detail how the agricultural revolution, because it changed how we interacted as humans, from being largely hunter, gatherers moving around to becoming sedentary. And how that just rewired our minds in many ways and how that has basically, resulted in some of the good and bad of the world we live in today. And I just think about social media and the kind of digitization of the world, as kind of like a modern form of the agricultural revolution. It's something that's just really transforming the pathways of human connection. It's transforming the ways people engage. It's transforming the way people move and talk to each other.

Fadi Quran: And I think we are not even aware of 5% of the actual consequences that this is going to have on their future. How is U.S. politics going to look like 20 years from now where this generation is just saturated by this type of content. It's terrifying to think about that because it's so unpredictable. But the small amount of data we already have now does not show a pretty picture.

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 20: The Bully's Pulpit

Fadi Quran: And I do feel, based on the meetings when we brought the survivors into the different offices in Silicon Valley, based on a lot of the, just communities we work with around the world. From high level politicians to indigenous communities across Africa and South Asia and so forth, I do believe that there is hope in creating a better world, but I think that hope is quickly being diminished by this social media environment that we have today. And it's on all of us. For everything that we care about to fight back against this and stop it before it's too late.

Tristan Harris: Yeah, I couldn't agree more, when I think about the impact on the next generation. Everything I just heard from this psychology professor about the quality of students' thinking and the cognition and the conversation with the U.S. senator saying they're answering to more and more crazy and extreme constituencies. There's just less high-quality conversation. Less nuanced thinking and that is debasing whether democracy has a future. It's a form of government that is based on a strong, sovereign mind. A powerful mind. Critical minds and conversational and open-minded minds.

Tristan Harris: And I think about a broader version of your "Correct the Record Campaign" because if you think about how this gets adjudicated. You and I and everybody listening to this, no-one wants this future. I mean, I don't think anybody listens to this and says, "You know what, Tristan and Fadi? I actually disagree. I think we should just keep plowing right straight along doing exactly what we're doing." I mean that's what give me kind of optimism is when everybody sees this, we all realize that no-one wants this.

Tristan Harris: And yet the question is, how do you reverse out of what is, in my mind, a 10-year long hypnotic trance of artificial polarization, of artificial loneliness and conditioning of children's minds in a certain way? I almost feel like we need a kind of truth and reconciliation commission to sort of say, back to the kind of "Correct the Record" thing. It's also "Correct Humanity".

Tristan Harris: We have to go back and say, "We the technology companies, in the same way that we would notify you about a fake news story and say this wasn't real, we have to notify you about a false consciousness that we've created that also wasn't real." And here specifically the effects that that has created. It's created some false loneliness, some false addiction, some false polarization, false extremism, losing a reputation of mainstream media because we actually had this role in turning mainstream media into click bait and exaggeration.

Tristan Harris: And I'm not trying to say that evil people in tech did this intentionally. I'm saying we allowed a bad business model to take over that none of us really want now that we see it. And we need to be able to not just correct the record but I think kind of have this correcting humanity. Correcting the false consciousness with this back notification to all of humanity. I mean I can imagine the day I wake up in the future where the close to three billion people collectively using Facebook, YouTube, Twitter et cetera, get not just a notification that says, "I read this false news article," but, "I've been inside of and marinating in a false consciousness. That technology created through a business model." And I hope that policy-makers listening to this will think about...

Tristan Harris: We all want this to change and so a path that would be a solution that's big enough, that would correct enough of this, would involve, I think, really naming the perimeter and surface area of the false consciousness that has taken place and reversing as much of it

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 20: The Bully's Pulpit

as we can. Because we want a future where sovereign minds and democracy actually works. We want a future where children aren't hollowed out and dumbed down and made more lonely and bullied more often. But instead are more free, more developed, have more inter-generational relationships and wisdom available to them.

Tristan Harris: And technology could do all of those things if we actually flip it around in a radical enough way. And that's, I think, with both of our work, how do we have campaigns that don't just try to stop the bleeding as it gushes out in more and more places, death by a thousand cuts. But how do we actually transform the solution to just mass healing from what's taken place. And I'm just curious if you have any thoughts on that or other things that you wish policy-makers or others would do to help us get back to a stable and sane place again?

Fadi Quran: First I have to say that I am deeply moved by that vision, Tristan. Look, it gave me goosebumps as you spoke about truth and reconciliation commission and everyone getting a notification saying you know, "Wake up from this false consciousness." That's a beautiful vision and I think it's also realistic. I think it's something that we can achieve together, so thank you for sharing that.

Fadi Quran: And on the question of policy-makers, definitely. We spoke a lot about our advocacy and Avaaz's work targeting the platforms. But the truth is, we hope that they act. We give the fact that they actually do act a less than 50% chance and that's why a big portion of our advocacy is now targeting policy-makers and decision-makers. To create smart, deliberative regulations that allow for more transparency that can hold the platforms accountable for the type of content that they are amplifying. And really create an actual living, breathing type of truth and reconciliation committee that starts fixing these problems one policy at a time.

Fadi Quran: And I think we speak about this at a pivotal moment, probably a tipping point where we have a small window of opportunity. And that's connected to the U.S. 2020 elections. And Avaaz, we're a bipartisan organization but we are working and investing a lot of our time to ensure that these elections are decided by democratic means, not by disinformation, not by interference. By safe means where people go to the polls armed with facts, not armed with anger and outrage and fake news.

Fadi Quran: And that is an important step because if that happens, if democracy actually wins in these elections then it will open the door for creating the type of policies. Whether that means instituting a transparency parameter that allows for detoxing the algorithm or makes it necessary that platforms for viral disinformation content provide a correction for the record. There's certain countries we feel are moving in that direction. The EU as a region is now considering how to regulate disinformation very seriously and we have hope that they'll move towards these solutions.

Fadi Quran: But I think, when we look at elections now as they're unfolding in the U.S. and because I think a lot of our audience are policy-makers in the U.S., we found that the top 100 fake news stories, viral fake news stories, that have been independently fact-checked by AP, PolitiFact and others, had reached, and this is from November, 155 million views on Facebook. So that's more than the registered voting population in the U.S. and that's 1.5 times more than the reach of the top fake news stories that were being spread six months ahead of the 2016 elections.

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 20: The Bully's Pulpit

Fadi Quran: So if you just look at those numbers, again, just look at the data, it looks like even these elections, with all the talk about fighting disinformation, with all the promises of taking down fake bots and so forth, we're looking at a tsunami that's about to hit the U.S. And we're beginning to see the first waves of it already. And so if we do want to create this vision, if we do want to fix this problem for the future we've created, every policy-maker listening to this, every platform executive needs to begin using their microphone. Using the power that was given to them by the voters who put them in congress or on the senate. To say, "We want to defend the fabric of the United States. We want people to go to the polls armed with facts." And the platforms need to defend that.

Fadi Quran: Because right now we're seeing a shift in the complete opposite direction and if that happens, you know these platforms, although the EU and other countries may move to regulate, these platforms are based in the USA and if the U.S. does not act to solve them and regulate them, it will be very hard to hold them accountable. But at least now, to ensure a path to smart regulation, policy-makers need to begin speaking up and saying enough is enough. And I don't know what you think about... I mean, I'd love to hear your thoughts on what you think needs to be done to push for more smart policy, particularly around the U.S.

Tristan Harris: Yeah, it sort of feels like the global climate change of culture. You know we often say with climate change, it's an enormous problem that would involve thousands of companies and industries and changing from the cement that they make to the agricultural methods to regenerative et cetera. All this stuff and we'd have to have hundreds of governments help enact the policy in their carbon taxes to try to transition that entire ecosystem to something that is not causing an ecological crisis.

Tristan Harris: And that can be really debilitating. And we can say, much like climate change this problem of sort of the mass downgrading and degrading of humans through this extractive business model of automating and manipulating human attention with AI at scale through these major platforms, can feel really hopeless. But unlike climate change, only about a handful of, I don't know, 100 to 200 people at a handful of companies, maybe 10 companies, regulated by one government in one country, the United States, could actually completely turn this around. If we had the political will.

Tristan Harris: Like if you think about how tractable this is to change these technology products compared to other issues where you have to change, again, hundreds of industries, hundreds of countries. Changing and regulating. This is immensely tractable. The only thing that's missing is the kind of political will and motivation that says, "Enough is enough." That we can no longer have this business model. That profits off of the kind of long-tail automation of user-generated content with automated algorithms choosing what goes in front of which eyeballs and cannot distinguish between good faith versus nefarious actors.

Tristan Harris: And so, you know where I'm optimistic is, if the United States or specifically California, as a state legislator, could actually act. And I think that pressure from all sides is helpful. The UK has actually got some stuff that's moving with the Ofcom Internet Harms Report that's moving along the European Union. The new president of the European Commission, Ursula von der Leyen, is laying out a new EU Digital Strategy which we'd like to see move to address and name some of these issues.

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 20: The Bully's Pulpit

Tristan Harris: But really, why not do it right here where the companies are located and where we have the most power and influence? And I think what we have to do is create a vision that everybody loses in this world if we keep letting it go. It's not that one political side wins and the other side loses and ah ha there we go. We can dominate and then carry out the future. It's actually that this win-lose game being played over the minds of everyone, and with the most powerful country on earth here in the United States with nuclear weapons, is actually debasing the viability of that country. And I would say that the U.S. government is already a puppet or shadow government of the forces of polarization/disinformation that have already been sown by these companies.

Tristan Harris: And what we really need is this kind of, I think a policy that again, meets the scale of the problem. Something like not just "Correcting the Record" but "Correcting Humanity". And having that kind of mass truth and reconciliation for what has happened. Again, not to angrily blame evil people at tech companies, because that's not the case, but to actually say, "None of us wanted this to happen with this business model and no-one profits or wins from the next generation of children being harmed in the ways that they're harmed by bullying, depression, teen suicide increases, all of those kinds of things."

Tristan Harris: So, I think the future of our lives and of humanity really depends on, and actually on not even very complicated set of actions but just simply having the political will, that everybody loses if we don't act. And I think when you paint a picture of omni-lose-lose, that everybody loses by continuing on with this business model, that's the kind of representation we all need to see is down the road if we don't change.

Tristan Harris: It's not a matter of Fadi and Tristan want our way and other people want the other way. It's about the viability of human civilization and I hope that through your amazing advocacy, and people should really check out the petitions and campaigns that you're running on, is it [avaaz.org](http://avaaz.org)? That people can back and support the work that you're doing and that we continue working together to make this change happen, because each one of us has a role in making it happen.

Tristan Harris: And I just couldn't be more grateful for the fact that there's not that many people out there working on this. There's not that many people who have spoken up about it and there's certainly almost no-one that I know who's actually done the specific work you've done by bringing the people closest to the pain closest to the power. And so I just want to thank you so much for coming on the podcast and for sharing all of your insights and knowledge with us today.

Fadi Quran: I want to also thank you and the team at the Center for Humane Technology, honestly, for a lot of the advice and a lot of the key talking points. But I also want to highlight that a lot of the terms that have been introduced to this conversation, such as freedom of speech is not freedom of reach, are things that you and your team helped develop and spread. And that's already shifted this conversation in the right direction. And so we'll definitely continue working together until this problem is solved and I'm excited to see the future that we create, where the technology that we have adds to humanity and doesn't subtract from it.

Tristan Harris: Me too. And thank you so much, Fadi, it means a lot.

**Center for Humane Technology | Your Undivided Attention Podcast**  
**Episode 20: The Bully's Pulpit**

Tristan Harris: 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.