Tristan Harris: Imagine if US politics worked like this today.

Dick Gephardt: I used to put 60 members in a room every night at 5:00, 60 Democrats, to talk about the issues that we had to figure out. I had people from far left to far right, everything in between, all kinds of views, all kinds of disagreements on every issue. That's how the discussion started. I used to say to members, "None of us knows everything. We have to listen to one another. We have to learn from one another." After we had had 20 meetings, people would come up to me and say, "When we started this, I just knew what our policy should be. I listened to him and her and I changed my mind." That's human interaction. We can come to agreement, compromise, consensus on how to solve controversial problems.

Tristan Harris: That's former congressman and House Majority Leader Dick Gephardt, who served in the US Congress for 28 years. He's referring to a time just two decades ago that seems so far away now with how hard it is to imagine American elected officials seeing eye to eye. I'm Tristan Harris, and this is Your Undivided Attention, the podcast from the Center for Humane Technology. If that's what it took to build consensus in Congress 20 years ago when things were far less polarized, imagine what kind of things it'll take to build consensus now. But when it comes to social media risk, there is reason to hope for consensus.

Because in a way, we should be able to agree on the fairy thing that's preventing us from agreeing on anything. Division is a unifying issue, as we just said on 60 Minutes. Back in October of this year, I went to the launch of a new initiative called the Council for Responsible Social Media in Washington, DC and Congressman Gephardt was there. It left me feeling really excited because this was a coalition of public health experts, national security leaders, religious leaders, people who study these issues, former representatives and congress from both sides, people who just care about making our democracy work.

The changes that will come out of this will make a major difference for users of these platforms around the world, not just in the United States. During the event, I sat down with Facebook whistleblower Frances Haugen, who obviously many of you know is a friend of the Center for Humane Technology, and we wanted to release that conversation today on Your Undivided Attention because Frances and I really got to unpack some of the issues we both care about in a timely way and we think it's very relevant as we go into 2023. With that, here we go.

Speaker 1: I'm very honored to introduce Tristan Harris and Frances Haugen.

Tristan Harris: Can you all hear us? Yes.

Frances Haugen: Can you hear us? I was a little worried about the people in the back.
Tristan Harris: Francs, let's start with you. I think people underestimate the amount of risk that you took on to come out. I mean, you risk your life going to prison, having a trillion dollar company throw bazillions of dollars in legal fees at you. Could you say why you were willing to take that risk?

Frances Haugen: I was working on how was Facebook influencing the information environment on some of the most vulnerable places in the world. Most people are not aware that Facebook went into countries around Africa, around South America, around Southeast Asia and said, "Use our products. The internet is free." Facebook became the internet for at least a billion or a billion and a half people. The version of Facebook that we use, of Instagram that we use is the safest, cleanest version of Facebook in the world.

I genuinely, genuinely fear that there are tens of millions of lives on the line in places like African countries and Southeast Asia. We've seen two genocides so far, Myanmar and Ethiopia. I think those are opening chapters of a dystopian novel that I desperately don't want to read the end of.

Tristan Harris: In talking to you, I know that you work with the Civic Integrity team. You would be looking at the 10 most popular posts in Ethiopia and seeing those beheading videos. It was things like that. Just give people a taste of the kinds of things that you were seeing. Because I think you open up a Facebook feed here. You say, "Okay, occasionally I see some political content. I see some weird divisive conspiracy theory going viral," but what were you seeing in those countries?

Frances Haugen: Facebook had a team that was called the Social Cohesion Team, which was the team that was responsible for fighting ethnic violence. We had a thing called virality review every week or every other week where they bring in translators and explain to us the most popular 10 posts in each country that were at risk, our tier one at risk countries. Every single post would be a beheading video. It'd be an accusation that the opposition was molesting children, that this child was kidnapped and tortured, like truly horrific misinformation. There's this question of why was this the most popular content in those countries when we don't see that here?

Facebook was spending 87% of its operational budget for misinformation on English, even though only 9% of users spoke English. I want to remind you, the United States is not a country of only English speakers. It is a country of many, many Spanish speakers, many, many other languages. They are not going to use as safe a version of Facebook as English speakers' Facebook. They wanted us not to ask questions, and they wanted us not to demand more. And that's why they hid the reality from us.

Tristan Harris: One of the things that's so interesting about the perception of you publicly is that you're known as the Midwesterner that cares about the kids. What you and I first started talking about when we got to know each other was actually these
global harms, breakdowns of democracy and ethnic violence, but you're known for the teen mental health harms. What do you think is the reason why that's true?

Frances Haugen: They're interconnected. The reason why we see the most viral content in those countries being these horrific pieces of content is that the algorithms push us towards extremism. That when you have algorithms that say content is better if it gets more clicks, it ignores the fact that the shortest path to a click is hate, anger, division. When it comes to kids, you can start with a blank Instagram account, no friends, no interests, and just do some very innocuous searches like healthy eating, healthy recipes.

Just by clicking on the first five pieces of content each day, within a couple of weeks being taken to pro-anorexia content. If you're mildly depressed, it can push you towards self-harm content. These algorithm escalate. It's an escalator that goes up ever more.

Tristan Harris: In other words, as you said, the harms are interconnected. What seems like we have one set of harms for teen mental health happening over here, so we should bite that off and chew it. Let's solve that problem. But then we have this other set of harms called ethnic violence or breakdown of democracy or conspiracy theories. Let's bite that problem off and chew it. But what you're saying is they're interconnected. I think what we should do now is just quickly talk about what are the drivers for how all these problems are going on.

Just to say a little bit of our perspective, starting in 2013 when I was setting the attention economy at Google, it was pretty clear that there were these key drivers that were going to drive the world culturally in specific directions. Because if you have a certain set of base facts, okay, we have a private company worth a trillion dollars whose business model is they have to get more attention every year.

Frances Haugen: Their business model is ads. The more attention you give, the more ads they give.

Tristan Harris: Each of these companies has an EGO or an embedded growth obligation. They have to grow every year and every quarter. With that EGO also comes, it's growing what? Do they grow just data that they could collect? Well, yes, but they also have to grow the amount of attention that they get from humanity. If I don't get that attention, the other one will. If Instagram doesn't add a beautification filter to match TikTok in the arms race for teenagers' mental health, Instagram's just going to lose the arms race. It's pretty simple game theory. But when you then say, "Okay, if I don't do the three second videos versus the 30 second videos, I'm going to lose to the guy that does the three second videos."
When you play that out, this race for attention, starting in 2013, the reason that I came out, my version of Frances’ story, is that we can predict the future. I can tell you exactly what society’s going to look like if you let this race continue: population centric information warfare, weakening teenage mental health, shortening attention spans, beautification filters, unrealistic standards of beauty for teenagers, more polarization extreme content, more conspiracy theories. These are all predictable phrases that describe a future that if you allow this to continue, I can tell you exactly what the world’s going to look like.

Part of the reason we're all here today is to not just talk about those problems, we want to solve them, because we know that this leads to a total dystopian catastrophe novel that unfortunately is playing out true every day.

Frances Haugen:

I want to unpack that a little bit. We've heard people say things like they're intentionally designing these systems for anger. They're intentionally designing them for division. One of the things that I was really struck by when I went to Facebook was how kind and conscientious the people were that work there. The kind of people who work at social media companies are people who value connection, right? They're not shadowy figures.

But what Tristan's talking about here about the market incentives, the fact that these are private companies that we are asking to run critical public infrastructure in a completely untransparent way, we're asking them to maintain public safety, to maintain national security when those are cost centers, they're not profit centers. You end up in a situation where they want to do better, but because they have to meet these market incentives each year, it's hard for them to get there. I guess the question I have for you Tristan is like, what conversation should we be having then?

Tristan Harris:

In preparation for answering that question, I think one thing we have to notice is that per the E. O. Wilson quote that we always go back to, that the fundamental problem of humanity is we have Paleolithic emotions and brains, medieval institutions, and accelerating godlike technology. I repeat it thousands and thousands of times because of how true it is and how deep it is as an insight to then see how do we solve a problem. Part of the medieval institutions is that law always lags the new tech. We don't need a new conception of privacy until you have ubiquitous cameras that start getting rolled out in the 1900s.

We don't need a right to be forgotten until new 21st century technology can remember you forever. One of the problems is that we have technology moving so fast in the current regulatory environment is, okay, well, I have these existing moral philosophies of privacy and data protection. And these are good. We want these things. But notice that the breakdown of teenage mental health or extremism in Ethiopia or this arms race for attention and engagement, it's an adjacent and slightly different set of areas and we don't have laws or moral conceptions for those areas.
Frances Haugen: Often when we write laws, we write them about externalities, right? That when we have the system operating in isolation, there are incentives where these...

Tristan Harris: What is an externality?

Frances Haugen: An externality is when there's a cost. Let's say Facebook's going ahead. They're getting you to pay attention. They're getting you to click on ads. They get money for those ads. They're offloading onto you though the anxiety that's building in your heart, the child that took their own life, the political division at Thanksgiving.

Tristan Harris: Those don't show up on the balance sheet of Facebook. They don't have to deal with the Thanksgiving conversations that don't work anymore.

Frances Haugen: The thing we want to emphasize is that there are really good, really simple practical solutions that would reduce a lot of these problems. It's things like that escalator where you keep going for more and more extreme content. When we talk to pediatricians, when we talk to child psychologists, they say, "Kids get that this is happening." They get that when they go on there, they feel more anxious. They get that it's making their eating disorder worse, but they're being forced to choose between their past and their future. They can give up their account, but they have to give up all their friends and their connections.

They have to give up all their past memories. Kids aren't willing to give up their past for their future. They should be allowed to reset the model any time they want to. Any of you should be allowed to reset your model. You should have that right, even if it's going to make Facebook less money. It's things like saying, how do you put mindfulness in the sharing process? Do you require people to click a link before you share it? Or things as simple as, what level of hypervirility do we want to endorse?

When something gets beyond friends of friends, imagine a world where instead of having a little reshare button where you can keep spreading the misinformation we said, "We value choice. We value intentionality. You can say whatever you want, but once it gets beyond friend of friends, you have to copy and paste if you want to spread it further." That change sounds pedantic. You're like, "Frances, why are you asking me about colors on share boxes or share fingers?" The reality is that simple change has the same impact on misinformation as the entire third party fact checking program. Only no individual is now saying this is a good idea or a bad idea.

Tristan Harris: Let's actually break that down because this is a profound statement that we were both... You said it first. It came out of Frances' disclosures to The Wall Street Journal. That in Facebook's own research, simply taking away the share button and having you say, "I can still copy and paste the text manually and
Frances Haugen: I have to intentionally do it.

Tristan Harris: I have to intentionally do it.

Frances Haugen: Not mindlessly.

Tristan Harris: We're talking about a tiny change, something that a JavaScript engineer can spend a day and it's done. That would be more effective than, I think you said in the documents, a billion dollars spent on content moderation and all the other sort of trust and safety initiatives.

Frances Haugen: I don't know about all that, but the third party fact checking program where they pay journalists to go out there and write articles and say, "This link, this concept is no longer allowed on our platform."

Tristan Harris: I think this gets to the point then. Why wouldn't a trivial change that an engineer could make in one day, why isn't that happening?

Frances Haugen: This comes back to this question around externalities and incentives. The reason why we have to push for things like platform transparency, so the PATA Act, Platform Transparency and Accountability Act, it would allow us to see inside those companies. What would it look like in terms of what people would be willing to stand up and demand if they could see for themselves that data instead of just taking my word for it or looking at the documents that I brought out?

Tristan Harris: The Platform Accountability and Transparency Act, or PATA, has not yet been introduced to Congress, but it's a bill with bipartisan support that obliges tech companies like Facebook to open their data to researchers so that we can actually study the effects of these platforms in a meaningful way. The bill came about in direct response to disclosures from Frances and other social media whistleblowers. We don't want to live in a world where we have to wait for the next Frances Haugen or the next whistleblower to know what's going on inside these platforms.

Frances Haugen: We have to have the ability to have those countervailing incentives, because otherwise the profit motive will just keep pushing away from these really simple sensible changes. The great irony, and I've had to repeat this in every interview since then, one of the core parts of my testimony was the idea that when we focus on content moderation, solving these problems after the fact, it doesn't just distract us from real solutions, it leaves behind everyone who doesn't speak one of the 20 biggest languages in the world. That's what causes that ethnic violence that I talked about in the beginning.
Tristan Harris: One of the things that I find so ironic is when Frances came out, I actually was in a gym in New York watching from a hotel room and I was on the elliptical and I was working out. I was watching live her give the Senate testimony hearing. I started crying because I was so moved by the fact that she had done something that I had never seen happen, which was she got bipartisan consensus. You had senators from the left and the right saying, "We don't disagree. We're going to solve this problem."

And then the next day, stories went viral on Facebook and Twitter saying that Frances was some kind of operative, or that she was actually a plant by the government to justify more censorship, or that Facebook planted... It was like more and more bizarre stuff. She had a meeting next day with one of the major senators who the day before had agreed and said, "We got to do something about this." I won't say the name, but this is someone who doesn't typically agree on these issues. The senator canceled the meeting.

One of the reasons I wanted to bring this story up is that when people so distrust your motivations as a person, the point is that our society runs on trust and Facebook and Twitter are trust degrading machines because they reward those who are more and more innovative and coming up with new species of cynicism and distrust. You will be paid 100x more likes, followers, rewards, and reach than if you do not do that. When you set a society up like that, I think one of the things that's actually helpful about that is we've been living in this 24/7... I was a magician as a kid.

We've been living in this 24/7 magic trick that everyone's cynical about everything, when really there's a handful of people who are cynical and we gave them 100x reach. There actually are trivial solutions like the things that Frances was talking about. She just mentioned one. There's a hundred things that we could do and we could do them tomorrow.

Frances Haugen: And we agree on them.

Tristan Harris: And we agree on them. The disagreement is an illusion. It's a magic trick.

Frances Haugen: I think we're running out of time, so we'll hit on just one or two more quick things. First, national security. One of the big themes for this council is on national security. People need to understand, what we're facing now is a national security problem. We had a major whistleblower come out in the last couple months. Peiter Zatko, I always mispronounce his name, he came out and he said, "Twitter is running on duct tape and gum. Their data centers are running hundreds of thousands of computers on out of date software because they can't staff the people to update them. But if you took out one or two data centers at certain times, all of Twitter could have been lost permanently."
Here's a clip from Twitter whistleblower Peiter Zatko's testimony before the Senate Judiciary Committee this past September, and we'll include a link to the whole thing on our show notes.

I'm here today because Twitter leadership is misleading the public, lawmakers, regulators, and even its own board of directors. What I discovered when I joined Twitter was that this enormously influential company was over a decade behind industry security standards. The company's cybersecurity failures make it vulnerable to exploitation, causing real harm to real people. When an influential media platform can be compromised by teenagers, thieves, and spies and the company repeatedly creates security problems on their own, this is a big deal for all of us.

When I saw this testimony, a giant red light went off for me, because the thing that we don't understand is that right now we're going to Mars with SpaceX because there is a pipeline, a supply chain of aeronautical engineers and PhDs. Every security professional that worked at Twitter learned it on the job. They learned it at Twitter. They learned it at Facebook and went to Twitter. They learned at a handful of other companies, and Twitter is not profitable or sexy enough to get enough really basic security people. We have a huge supply chain problem and it is endangering our national security.

One of the things you told me and then I want to switch to optimism is when you were working at Facebook, you left I think after the 2020 election, and you said, "Up until that point, we had these teams that were staffing election integrity, election safety," and then you said, "After the election, everyone was so burnt out. It was like we lost the senior neuro technicians for the brain implant and now we got the team who's left to try to keep handling our systems." Because of this talent issue you're talking about, people also get burnt out. They're aware of the problems that we're talking about.

There's this kind of moral internal issue where it's very hard to actually get people who can solve these problems, especially through the growing success of this kind of movement that's revealing that there's a problem. We do need to have more moral courage and more people working at these companies. That's why Frances said at the beginning, "I want more talented people working at Facebook, not less as a result of me coming out."

One of the reasons why the disclosures is so large, it's 22,000 pages, I literally wore a back brace at points because hunched over my computer to take these photos. It was that hard. The reason I brought them out was that top tier, the only reason why they were able to ask the questions they asked, to have the intuition they had was because a lot of these people were veterans that had been at Facebook for seven, eight years and they were all retiring. They dissolved Civic Integrity and something like two-thirds of our team just left the
company. A lot of those people didn't go into security in other companies, they went into other things because they couldn't do it anymore.

Tristan Harris: Of course, Facebook's recent announcement that it's laying off 11,000 of its employees will almost certainly push the company in the opposite direction, less security and more risk, because they have to let go of many of the people who work on the trusted integrity teams that exist to help keep the platform safe.

Frances Haugen: We have military academies because we understand that national security is so complicated and nuanced that we have to train leaders who understand the complexity of how to keep the nation safe.

Tristan Harris: I think these problems can feel very overwhelming. They're massive.

Frances Haugen: We need scaffolding to fix it.

Tristan Harris: We need a scaffolding model of our minds. How are we going to build up towards those solutions? There's triage solutions which are things that existing companies can do today sometimes in the span of a day, an hour or a week or a month, that when, for example, Apple makes changes to the privacy features, that can take $10 billion a year off of Facebook's revenue and the entire surveillance capitalism kind of model by Apple being a short-term regulator and interim solution.

Frances Haugen: There's solutions on teen mental health where we could turn them on in the next two weeks and you probably cut the suicide rate the next year by 25%. That sounds ridiculous, but I could put a whiteboard up here and just lay it out for you.

Tristan Harris: There's triage solutions and we should do all those and we should make lists of them. And frankly, we shouldn't have just a handful of people in this room. We should have a collective intelligence, almost like an incentive prize. We want as many people in the world saying, within the boundaries of the current incentive model, what are the immediate things that would make the biggest difference? That for the sake of getting more good faith returns from the public, what can these companies do? Then there's transition, which is, how do we actually set the turns in motion to catalyze a shift in the incentives and the structure?

Things like passing the Platform Accountability and Transparency Act. Just like we have cybersecurity requirements now growing on companies, can we have different requirements on national security from government? And then we also have long-term solutions, which is that in the long-term category, I think what we need to do is get from a thinking model that's, how do we get to less toxic social media to instead tech plus democracy equals stronger democracy? Because right now China is employing, the CCP is employing the full suite of exponential technologies to make a new form of 21st century authoritarianism.
When we look at democracies, we're not employing the full suite of 21st century technologies to make new and upgraded forms of democracy. We're instead allowing private business models to profit from the degradation of democracies. The long-term has to be, "What is a vision for a digital democratic future?"

Frances Haugen:

To end on an optimistic note, because I talk about things like ethnic violence a lot and I have noticed it can be a downer. To end on an optimistic note, because we can't change if we don't have hope, I want to remind everyone, our democracy has gone through crisis before. I have only had the right to vote as a woman for about a hundred years. Contemplate that. That's insane. African Americans basically didn't have the right to vote in large parts of this country as recently as the '60s or even more recently. We have had crisis before. The thing about democracy is it's resilient. That democracy is a tree that grows because we tend to it, but it can heal, it can grow back.

When it comes to new forms of media, every single time before when we have invented a new form of media, it has been incredibly disruptive. It might get substantially worse before it gets better. But every single time before, we have figured out how to respond. When we invented the cheap printing press newspapers, we literally had wars over misinformation, but we invented journalism schools and journalistic ethics and transparency rules over who owns what and how much concentration can you have. When we invented the telegraph, this sounds crazy, the telegraph did not cause the Civil War, but historians believe it influenced when the Civil War was. Because suddenly people could see that they disagreed with Washington, that they disagreed with Maine, they disagreed with Minnesota. Suddenly, we had one nation and we had to work through those problems together. Radio. Radio gave us personal relationships with our leaders and it led to a lot of the movements that caused World War II. This has happened over and over and over again. The reason why it feels overwhelming right now, why it feels like we don't know how to get out of it is because this is our time to shine.

This is our moment to stand up and do something. We are going to figure out how to get out of this, but it takes work and we have to do it now. But I have faith that we're going to make it through this. Thank you.

Tristan Harris:

Let's solve it together. The Council for Responsible Social Media will be working to pass the Platform Accountability and Transparency Act in the next Congress, as well as social media safety regulations like the Kids Online Safety Act or KOSA. In its next campaigns, we'll advocate for other concrete actions that social media platforms can take to ensure that they're not undermining US National Security. To learn more about the Council for Responsible Social Media, which is a project of Issue One, you can follow the link in the show notes. We'll also post a link to the interview we did with Frances Haugen back in October 2021.
Your Undivided Attention is produced by the Center for Humane Technology, a nonprofit organization working to catalyze a humane future. Our senior producer is Julia Scott. Mixing on this episode by Jeff Sudakin. Original music and sound design by Ryan and Hayes Holiday, and a special thanks to the whole Center for Humane Technology team for making this podcast possible. You can find show notes, transcripts, and much more at humanetech.com.

A very special thanks to our generous lead supporters, including the Omidyar Network, Craig Newmark Philanthropies, and the Evolve Foundation, among many others. And if you made it all the way here, let me just give one more thank you to you for giving us your undivided attention.