DarkHorse Podcast with Tristan Harris & Bret Weinstein (192k...

Sat, 10/9 12:24PM • 1:46:19

**SUMMARY KEYWORDS**

point, problem, facebook, process, technology, people, attention, evolutionary, apple, shared, function, heuristic, fact, governed, asymmetry, reality, twitter, algorithms, question, world

**SPEAKERS**

Tristan Harris, Bret

**Bret** 00:05

Hey folks, welcome to the Dark Horse podcast, I have the distinct pleasure of sitting today with Tristan Harris, who is co founder of the Center for humane technology. And the host of the center's podcast, your undivided attention. Tristan, welcome to Dark Horse.

**Tristan Harris** 00:22

It's great to see red Good to be here with you.

**Bret** 00:24

Yes, it's great to be here with you, just so that people understand where we're coming from you and I know each other prior to either of our podcasts. We're friends. And so we've had a lot of informal, behind the scenes discussion about topics that I'm sure we will get to today. You are in the Bay Area. Am I correct?

**Tristan Harris** 00:44

Yep, in the San Francisco Bay Area, after a brief sojourn due to a climate outcome of our house burning down due to the wildfires in California in September. So issues of anyway, that's just the reality of my situation. And back in the Bay Area,

**Bret** 01:02

bay area that Yes, that sounds slightly beyond what we usually think of as a climate outcome. You know, the Earth is getting warmer, but it's rarely hot enough to catch your house on fire cheese, your house burned down. That is shocking.

**Tristan Harris** 01:16

Well, the crazy thing is that the we obviously don't have to go into this much. But the house burned down two weeks after the release of the social dilemma, which was this global sort of phenomenon. So actually, I had less than 24 hours between that happening, and then doing the today show at you know, five in the morning and continuing to do interviews and you know, everything that was happening because the film, the social dilemma was seen by something like 100 million people in 190 countries and in 30 languages. So it's it's, you know, balancing all that after the house burned down and being sort of on the road and having to work it all out was was just a unique experience to have.

**Bret** 01:56

Yeah, that's, that's remarkable. I should say, I was going to mention your amazing Netflix documentary, anybody who has not seen it should absolutely seek it out and watch it, it will tell you a lot about what's been on your mind and why I thought it did a very good job of putting flesh on the bones that are sometimes hard to get people to, to see where, you know, everyone has a sort of vague sense of how big tech works and what influence that may be having on things like our collective sense making. But very few people have any idea of what the mechanistic underpinnings would be. So the documentary did a good job of, of outlining that you want to say anything about what the what the core message of the documentary is?

**Tristan Harris** 02:53

Yeah, I mean, you know, the film is really a message about the people who are inside the technology companies, you know, Justin Rosenstein, who is co inventor of the like button at Facebook, and the person who brought the business model of advertising into Facebook and early employees at Twitter, at Facebook, at YouTube, etc, saying, you know, this business model is fundamentally misaligned with the fabric of society being healthy. And the, you know, it does that by by explaining how the same problems that are with the technology companies are really part of our economic system, that a economic logic that lives on an infinite growth paradigm on a finite substrate, usually it's the planet causing climate change. In this case, we have an infinite growth economic paradigm of big technology companies based on a finite substrate of human attention. And that, I think that the film is really summarized in Justin Rosenstein his line at the end that he says, so long as a whale is worth more dead as whale meat than alive. And a tree is worth more as two by fours, that is a tree in our attention economy. To technology companies, human beings are worth more as dead slabs of predictable human behavior than as living, breathing choice making meaning Mickey you know, you know, semi free willed humans, and that that logic is really the core of where you'll see lots of problems. So shortening of attention spans political polarization, things like this.

**Bret** 04:25

So of course, as an evolutionary biologist, I see all of these things and I find no and of analogies to my home territory, and I try not to drag people there more than necessary because if it's not your, your home court, it is, you know, it adds an extra kind of step to one's logic, but there's so much here. You know, I'm used to situations in which a very small advantage over enough time results in the modification of creatures and And what we have here with respect to the attention economy and big tech algorithms, etc, is not so subtle influences operating over a short time but with completely predictable results. And the answer is it doesn't really take very much. If the signal is consistent, it doesn't take very much to drag Billy, for example, our shared narrative, which is so necessary to us functioning as a society, it doesn't take much to drag it into some territory where it serves a private interest at the expense of the public. And, you know, I think if people realized how delicate the assumptions were on which normal functioning depends, they would be less shocked by how distorted things have become, in light of these novel influences that even if they were well intentioned, when they were originally constructed, are just simply incapable of maintaining the, the foundations of thought.

**Tristan Harris** 06:05

Yeah, yeah, the, I think, the breakdown of our shared reality due to a profit motive in which personalization is profitable. This is really important and subtle to get I mean, it's sort of obvious once you see it, but giving each person their own Truman Show reality is more profitable and better at keeping your attention glued, than to creating shared realities that are not personalized, a personalized Twitter with a personalized algorithmically curated newsfeed is going to outcompete a chronological Twitter that just shows you tweets showing up on you know, whatever level they do. So it really becomes this race to hyper personalized, which then this race to narrow each of our individual construction of reality. And I think that is actually the deepest harm that I like to focus on these days. Because that's really what we're seeing. I think, in January 6, we're seeing it in the country, we're seeing it around the world, we do not have a shared basis of ground reality. I know you talked to our mutual friend, Daniel schmuck and Berger about some of these topics. But I think that is the kind of shared problem statement that we're now left with, I sometimes think of it bred to sort of play with metaphors that you are and Heather are talking about, that we now have a few different problems, we have a technological problem of how social media is business models, and fundamental design and morality and the kind of hijacking of all of our different evolutionary impulses, causes all these sort of, you know, cultural Fallout, this kind of cultural externalities, of shortening of attention spans, mental health problems of kids, etc. And this breakdown of reality, but now we have a cultural set of problems. The problem is metastasize. So now if you took technology away, our minds are all running, you know, not that they were never not running malware, but they're running much worse malware than they would have ever run before. And it's been exacerbated. So I think of it kind of like we had the Zuckerberg Institute of virology. And that Institute was doing Ghanem function research on mimetics on what kinds of mimetic viruses would work and they were literally having engineers tweak things like you know, the newsfeed recommendations, you know, what kind of things would keep people clicking more and they started evolving, hyper evolving in a certain pathway. In the same way they gain a function, I guess, research kind of works. They were also tweaking the friend recommendation systems to say Oh, which friends could we recommend for you to add that would cause you to stay on the product longer and more friends like you more like minded friends, more coupon friends, if you start off in human on more, work astan friends, if you start off in Pakistan, more magazine, friends, if you start off in Mega Stan, that will increase people's sort of tribal affiliations, then they were doing gainer function research on Facebook groups, what groups could we recommend you to join? And we now know from Facebook's own internal documents that 64% of the extremist groups that people joined, were due to Facebook's own recommendation systems. In other words, this was not a user who's going into Facebook and being like, what group? What's selling club? What book club? Do I want to join? You know what Dark Horse podcast discussion group do I want to join? instead? They're being recommended something and my friend Renee de resta, who's sort of a researcher on these topics, says it really perfectly when she was a new mom, she joined in organic baby food moms group, like where you make your own baby food, sort of a natural thing you might want to do. But what was the most recommended group? Do you think for the new organic baby food moms from Facebook's point of view? What would keep people what would people will keep users who looked like organic Do It Yourself baby food moms on the site the longest if you're just building a naive AI recommendation system,

**Bret** 09:32

I am frightened to guess for many reasons. I'm going to I'm going to resist guessing because I will tell you if I'm right here in a second you tell me what it was.

**Tristan Harris** 09:46

I mean, it's not maybe that surprising that it was more of the anti vaccine moms groups right because exactly

**Bret** 09:51

what I was gonna guess I didn't want to. I didn't want to say it because I didn't want to add fuel to the fire. Right? Right.

**Tristan Harris** 09:59

Yeah, well So, so and then once you joined that group, then Facebook said, oh well what do the anti vaccine moms groups, what do they tend to click on and what what keeps them around the most and recommended pizza gate and chemtrails and other things like that. Now when I say this, I might sound like a blue church left person who's saying, oh, all these conspiracy theories and all these sort of hesitations that people would have, are somehow naively informed, as opposed to there's a pre trans fallacy for each of these different sort of concepts. And I want to make sure that we're caveat in that. But when you understand that this gain of function research, Zuckerberg Institute of Neurology was doing this research that was hyper accelerating these kinds of evolved pathways. And now that that, that mimetic virus, that sort of reality dividing virus, that extreme defying reality, virus, left the Zuckerberg Institute of Neurology in Menlo Park, and actually escaped into the world. And it actually has shut down the global sense making apology. It is now just in the same way that COVID has shut down a global economy. I think that this mimetic virus has shut down our ability to see reality and shared waves. And I think that's really the problem is how do we not just dealing with the pandemic, but how do we deal with the kind of reality breakdown pandemic? And how do we have not just like reality bridging conversations? But really, what is the solution to that, and I think we're all trying to figure that

**Bret** 11:23

out. All right, I want to slow down a little bit, because it happens that you have landed, just like a millimeter away from the topic I was going to introduce here and get your thinking on it. So I know we're very close to the same page here. But one of the things that I think is most difficult to convey is that there is a natural synergistic relationship between human consciousness and creativity, and evolutionary dynamics that we are unaware taking place. And so I think we have, you know, there's a dichotomy, a false one, where people expect AI to take off and come for us. But they also intuitively know that that can't be where we are, because all of the devices that are supposed to be intelligent, are way stupider than they should be, right. So we sort of know that we've got some time before AI gets to that place. However, what we don't into it, is that humans can partner with semi artificially intelligent or narrowly artificially intelligent or even just learning algorithms. And they can do the part that the artificial entities can't do themselves, right. So you have a bunch of people inside Facebook, thinking about how to enhance Facebook's financial well being, right, just increase shareholder value at this point. And they can narrow the search space of possible functional strategies to a very small, very small canvas. And then things like a B testing can do the heavy lifting, which is, incidentally, you've landed on the perfect analogy, it is exactly again, a function research because what gain of function research does are not all of it. But the part of it that is called serial passaging involves taking a natural entity, or a largely natural entity or a composite of natural entities, and placing it in an environment so that evolution can do the stuff we don't yet know how to do. Right? If we could do it directly, we would, but we can't. So what we do is we just put it through an accelerated process of evolution, we basically reduce the noise and increase the signal. And then we take what comes out the other side, and we discover what we didn't know enough to desire. So that is exactly what will happen here. But imagine, you know, you could have had 1000 Facebook's right 1000, hopeful companies starting out trying to capture this space. And selection would have found the ones that happened to have the right things to keep attention on their site, right to direct people to stuff that would reinforce their, their, their biases, etc. Those companies would who would have risen to the surface instead, we had a couple of Facebook like entities that we had intelligent people within them, looking for mechanisms, you know, how does Facebook get ahead of my space, right? And that has been like a cyborg, where you have a partially undirected process of evolution and partially human creativity, fused together to discover defects in all of our cognition that allow them to steer us into behavior that we otherwise wouldn't be involved in. And, you know, it is. I think the thing is, if you understand that this isn't wholly dependent on an evolutionary process that is haphazard, nor is it wholly dependent on a conspiracy in which people simply sideline human well being and go for broke, but the two things can partner, it would be hard to overrate the danger of that. And what you've been saying forever, is effectively It's here. And we don't actually know the full range of hazards that emerged from it.

**Tristan Harris** 15:37

Yeah, that's it. That's exactly right. I'm so excited to have this conversation with you. Because I don't get to talk to evolutionary theorists where we can really go deep into each of the sub aspects of this problem. The first thing you talked about was this sort of projecting into the future, we often look into the future and say, when is AI going to take over humanity and we think about the singularity, you think about numbers like 2050 2060, you think about super intelligence, Nick Bostrom, you think about AI taking our jobs. But there's this much earlier point that you've made. I know ages ago, in your work, I remember you and Eric Weinstein, both your brother Eric talking about this, more like the accession point that technology doesn't actually overwhelm human strength that undermines human evolutionary weaknesses. And there's this sort of takeover point. And by the way, that point for me is more informed by my training and magic. Since the house burned out, I've been buying back all my sort of magic supplies, all this sort of these packages arriving with these kind of interesting magic goodies. An interesting thing about magic is it's based on an asymmetry of understanding, where the magician essentially does know something about an evolutionary weakness in your mind that you do not know about yourself. And it's that asymmetry of information that allows the illusion to work. And the problem is that technology has become this sort of 24 seven, hyper asymmetric magic trick, but we don't perceive it that way. Because we're enjoying so many things from it. We enjoy social validation and approval, we enjoy getting, you know, variable schedule, word slot machine type things, we, you know, we enjoy photo filters that make us look more attractive briefly. But these are all essentially, exploits maladaptive, depending on how you see it exploits on key evolutionary weaknesses. And it's really a race, where the phrase the race to the bottom of the brainstem, it's a race to the deepest evolutionary weaknesses, or as your brother has said, a race to create a backdoor to the human soul, to figure out and reverse engineer, whatever would we would call sort of a dimensionality of freewill to turn more and more of those dimensions into predictable outcomes, meaning less and less free, and more and more commoditized dead slabs of human behavior. It's really important to get this because I think if you go back and like you said, it's not a conscious intention by Facebook or Twitter, to try to ruin the world, or to profit greed at all costs. I think a lot of this was incremental decisions in an evolutionary environment where they're competing for attention. You know, I actually remember because i a lot of our insights about this work come from knowing that people who built these products, I was friends with the founders of Instagram, I was at South by Southwest in 2009, when Facebook was competing with Twitter for who's going to create the newsfeed and they're both competing for that chronological kind of instant newsfeed of people posting things really often and having it feel more alive and more rapid. And a newsfeed, it's giving you new things constantly is going to outcompete a news feed, that only has new things once every five hours, right, because there's a freshness to it. So in each case, and I hope we go more into this, that the race to capture attention led to all sorts of decisions being made that were based on the evolutionary fitness of attention grabbing machines, but not for the evolutionary fitness of a societal cohesion or societal intelligence. Because what I really worry about is that the net result of all of this is when you don't have a shared reality, or an ability to agree on reality, or you have such emotional grievances and resentments because you've seen infinite evidence of hypocrisy, and awful things the other side has done that you do not want to, you don't even want to have a conversation about what you could agree about and what you can do about any existential problem, whether it's climate change, etc, that have shrinking timelines. So when you kind of add this all together, the question I'm always asking is, what is an attention economy? What is a digital environment? What is the cyborg, like fusion of the man machine human machine interface that leads to a smarter, democratic, plural, you know, collective psyche, but is making better and better choices together? And I think because we have limited time on various problems, especially climate change. That's the real question that I kind of land on, but we probably should reverse engineer a little bit more of how we got here with the evolutionary hacking of tech.

**Bret** 19:47

Sweet. Now unfortunately, when I talk to, to you and few others, each, you know, little soliloquy opens up, you know, 1012 hours worth of topic. that we could cover and so there's a bunch of stuff I want to recover here, and I'm sure we're gonna miss some important things.

**Tristan Harris** 20:04

But let's Why don't we try to capture a lot of it though, because I think it's gonna be good to flesh out. Yeah, I

**Bret** 20:08

think it's really important stuff. One has to do with magic. And I wanted to point to something that you didn't exactly say, but I'm pretty sure you'll agree with the thing that's really fascinating about magic, it's not so surprising for reasons we can come back to, that human perception has blind spots and defects that can be exploited, right. But there's a big difference between running into a con artist who you don't know as a con artist who is attempting to exploit those things for their gain. And a magician, where you go, and the person says, Hey, I'm a magician, I do magic. And you know, damn, well, they don't do magic, right? But the point is, you, you are constantly at a good Magic Show, trying to figure out how it's being done and unable to write. And so the point is, the fact that you're aware of what's going on at Facebook, Twitter, etc, is no protection. And this is one of the things that I think is most telling is that the very people who constructed these things, the smart ones are aware that they are not in control, that when they are the users, they're as vulnerable as anybody else. And, you know, the fact that they very often have very strong limits for their own children's interaction with these things, tells us what we're dealing with, we're dealing with something so potent, that the awareness that it is happening is no protection.

**Tristan Harris** 21:38

And that's what's so dangerous, right? Because it means that there is no sort of safety place where we actually still have control because as you know, in the film, the social dilemma, Tim Kendall, who you know, is a father and has a couple of kids and talks about how he was promising himself that he wouldn't use Pinterest, he was the president of Pinterest, and he found himself in the pantry, you know, scrolling on Pinterest, ignoring his own children, he said, this is classic irony. I'm the one who's building Pinterest on a daily basis, it's like getting high on your own supply. He can't get off the thing, even though he's making commitments to himself, to not do it. And that's actually one of the parts of the film that I think people really latched on to. Either Raskin, or my co founder at the Center for humane technology talks about how he actually wrote himself software to, you actually will appreciate the intervention that he created to make himself less addicted to Twitter and to Reddit. And you know, the way he did it, Brett, is he actually created a random slowdown. So as you use Facebook, or Twitter or Reddit more during the day, it would actually add a variable random level slowed down how long it took for pages to load, if you say, Do you, you know, if we put a time limit and put a seatbelt on and say, just like Apple does, and screentime, you can't do this, you know, don't don't look at more of this, you know, you've hit your 15 minute limit, your brain just gets upset at that and says nothing to get out of my way. I'm gonna keep going anyway, you're not me, You're not the boss of me, right? That's, that's the same Paleolithic emotions that sort of come up in that context. But if you do a random slowdown, where it just gets slightly more frustrating, not a lot more friends, it doesn't take forever, it just gets slightly more frustrating. Over time, and you just your brain, eventually it gives your brain enough time to catch up with its impulses, your prefrontal cortex can catch up with your sort of emotional impulses. And you know, when you're making this metaphor of, you know, we're losing control, it's almost like the inmates are running the asylum, but really, the amygdala is running the asylum are kind of reptilian brain is running the asylum, because if no one has control, and no one's prefrontal cortex, gets to sit in the pilot seat of our sense, making a choice making and instead we're left more and more with the lower parts of ourselves, then that's the representation of where all of civilization goes, because when you zoom way out, and this is when I was a design ethicist at Google prior to doing this work, my biggest concern was that, you know, we have 3 billion human beings were jacked into this real, you know, instantiation of the matrix, you know, we say like, we actually did build the matrix, it doesn't look like it did in the movie. And at least in the movie version of the matrix, we had a shared matrix, whereas in this matrix we have we have our own different hyper personalized Truman Show matrix.

**Bret** 24:07

All right, awesome. This is this is fantastic. All right. Let's go to the next part of what you raised before. So we've got magic, which demonstrates that the fact that you're aware of of these defects in your own perception is really no protection from having them gamed, but you also have let's open up the parallel realm. Okay. The parallel realm, I would say is the world of optical illusions.

**Tristan Harris** 24:34

Yes, just thinking the gray the checkerboard right is where you're going to go with this.

**Bret** 24:40

I was going to go with the general category. So I would say that the world of optical illusions, which is now effectively a realm of academic study, is all in search of evidence of how we perceive things based on there may be Word categories in this, but at least two categories of phenomena. One, were some heuristic, and we'll come back to what heuristics are in a second. But some heuristic in our perceptual apparatus decides something is one way, or it's another, and a carefully presented set of data can be perfectly ambiguous between two possibilities. And therefore the brain, you know, to the extent that, you know, some prior bit of cognition, or the way the light falls, suggests one way, right? It sees it that way. And then if something changes, you may see it the other way. So you know, these sculptures, are they convex or they concave? The Necker cube, you know, are you looking which corner of the Necker cube were you looking at? There's those things and then there are other heuristics where, when something shows up, so you know, we have a frame rate in the eye, many people will claim we don't have a frame rate in the eye, because the frame rate isn't regular in the same way that a movie film would be, but, but we do clearly have a frame rate in the eye, which is why, for example, when a car with perfectly regular wheels drives into the parking lot, you may see them turn backwards, because they advanced not quite to the position of the other spoke, you know, in one click of the frame. But in any case, you've got all of these cases where something in nature, if your retina saw it at this position, and that position, it would be safe to conclude that it had moved, right. So now if we flash things in a particular way, the Mario brother appears to move through the scene. But in fact, all the Mario brothers are perfectly static. Anyway. The point is, that is exploiting a perceptual heuristic to cause the brain to conclude wrong things, right? That you can then test are wrong, like the checkerboard case that you show where the two squares look very different in hue, based on a gradation, but in fact, you can check that they're identical. So anyway, what I want to get at is, these heuristics are there for your well being. Right, in fact, one of the key to any rational perception, if you tried to perceive every pixel on your retina and process it, you just, you'd never get through a frame, right? It's just too much data, right? So you've got to break it down into chunks that are so reliable, that you just get a huge payoff for using a subset of the data to draw the full conclusion. And it's very rarely wrong. But the fact that these things are riddled throughout our perception means that a magician, or a person who is interested in creating optical illusions, or a corporation that is interested in keeping you on their site, all find them, right. In other words, they're littered through your perception for your well being, but they are exploited by others all too easily, because of their ubiquity and predictable illness. So the question is, without an agreement that hey, it's not cool to game other people's perception. I mean, maybe it is cool if it's a surprise party. And the point is, you're going to enjoy the surprise party when it's revealed to you so it's okay for me to you know, to take advantage of your trust briefly. But this is no

**Tristan Harris** 28:15

game is a good example of that with Michael Douglas. Say it again. The film The game is a good example of that with Michael Douglas, where you can have a entire false reality constructed for you that then the end of the movie is the curtain gets pulled off. However, you've been thrown through the wringer, you thought you almost died. And there's this question of what is an ethical sort of false reality or manipulation of someone else's experience? If you know one way to do as you're saying is you sort of reveal at the other end, there's often transformational experiences or cults or transformational workshops, which actually do more confrontational alternate reality creation by manipulating your experience. But then they say at the end, it was all for your benefit. And there's this, I'm saying this actually, because what my motivation for this work, right, which I don't often talk about, but when I was at Google as a design emphasis, the basis for this inquiry of how technology is impacting society has to do with what is an ethical, asymmetric, manipulative, race persuasive relationship, when something is influencing some of your evolutionary weaknesses in or outside of your influence. And when is that in a grounded ethical, humane relationship? Because increase the degree of asymmetry between what technology knows about you that you do not know about yourself is going to grow the number if you plan an AI your brain you think of the newsfeed is like a kind of a chess game where your brain is the chessboard, and you think you're going to scroll one more video or watch one more thing, and then you're done. And you think that's going to be it. But you have this asymmetric AI that's playing out a trillion simulations from all the little human voodoo doll animals that it's the human social animals that it seemed to use Facebook with the exact same usage patterns like it's seen 1000 other people watch those same cat videos or 1000 other people watch those, you know, Brett Weinstein, pods, podcasts and it seen what is the next thing that I could show you, that would get you to stay. And it knows way more about that than you are aware of it because it's literally ran more simulation. So the degree of asymmetry over time is going to grow. Which means that we have to actually have a relationship with something that knows more about us than we know about ourselves that is ethical or grounded or humane. That brings us to, we shouldn't escape the topic of events. I'm not trying to diverge us or I know we keep opening up threads. But you then ask what are other relationships where one party knows way more about our psychological weaknesses than another, you might think about a therapist who is licensed, who says, Hey, I'm going to know your deepest, you know, secrets, I'm going to know your deepest, you know, vulnerabilities, your biggest concerns, your, you know, your inquiries for yourself, your sexual identity, your relationship, how you feel about your romantic partners, they're going to know all that information, and they can't go use that to manipulate you. So you spend more time in a therapy room, or they can't use that to sell it to some advertiser so that someone can basically say, Oh, you have these kinds of anxieties, guess which political party would love to know those anxieties that I would want to target you. So once we see it in that way, this is both opening up an inquiry. But I think we can, we can leverage insights from different fields, whether it's psychotherapy, counseling, law, you know, a lawyer knows everything about your vulnerabilities and your case more than you know about it. And you have to trust that agent to act on your behalf. And so, gosh, there's so many things that we can open up here, but where you get to in technology is you simply cannot allow categorically a business model based on exploiting that asymmetric relationship. What I like about this framing is it's a kind of a one and done way to simply nail the business model and say, it's super clear, you would never have a therapist whose business model could be advertising to political parties to manipulate you based on what they hear, you know, in the therapy room, or, you know, a priest who's sitting in a confession booth who listens to 2 billion people's confessions, literally like all the things that they know about you, then they have imagined worse than that the priest has a supercomputer next to them, it's literally calculating patterns of all confessions they've ever heard. So they're actually making predictions for the person who's about to walk into the confession booth before they even say, the confession that's on their mind. They know what that confession is going to be. And they can sell that to another party. This is the kind of asymmetry of power we now have. And so Gosh, I know we're opening up so many different threads, because we can still go deep into the evolutionary stuff and the cognitive biases, and there's so much to say there. But I wanted to make sure

**Bret** 32:18

I brought that in. Yeah, fascinating. And you're right, so many interesting threads. But let's, let's see if we can nail down the point that you and I are converging on here, there is either a basis for trust. In the case of the therapists, their actual regulations, there's ethical regulation, ethical being about the rules of psychology, for example, that say for you know, just to take one example here, if therapist is not allowed to sleep with you, right, there's some sort of waiting period, after you've left therapy before your therapist could start a sexual relationship with you. Right, that's a protection, because this person, as you point out, knows all about you, and therefore would be in a position to manipulate you, right? And then financially speaking, there is an explicit contract that you have with this person. And so those protections, I don't know that they work, or they don't work. But I know that the point is to establish a basis for trust that you actually know what it's made of. Right? That's right. And then there's another kind of situation. So let's say that this is the flip side, where somebody is bound to you in a way that your trust isn't about the rules, but it's about the fact that they don't have an interest in using this against you. Right, so family members, for example, you know, Heather has up, don't know that we've ever explicitly even talked about it. But it's very clear to both of us that she has the right within limits to lie to me, for our collective well being, right. So for example, if you know, I'm up against a deadline, and I need to get something done. And you know, our collective well being is dependent on it, she may adjust what it says in the calendar about when it's actually due. Right. Now there's a limit to how well that works. Because obviously, I know that it's a possibility. But the basic point is like, I know, she's not going to use this to her advantage and my disadvantage, so he's

**Tristan Harris** 34:12

not making money, the more she lies to you. Right, right, still using it for your best interest. And there's actually a shared partnership of your being your own children that creates skin in the game. So that relationship is grounded in a real sort of, you know, sense.

**Bret** 34:25

Yeah, the the term we use for your shared fate, right? Yes, we have shared fate. The point is, you're going to if you're going to lie to me, you're going to do it honorably, right. And there's obviously ways in which that, you know, started up with some guy and started lying to me that wouldn't be it would be the exact opposite. But in the absence of such a thing, she has licensed to figure out where the line is, and even if we slightly disagree over it, it's done with good intentions. And I should say this is now uncomfortable, but I also find that my relationship with my children involved A huge amount of my gaming their perceptual apparatus. My kids are, you know, 16 and 14 at this point, but this has been true all along. And it is always done in the context of humor. But the basic point is some part of me knows that they live in a world in which they're going to be confronted with bullshit, right? every waking hour, right? There's going to be advertisements, conmen algorithms, all sorts of stuff. And so I use my the license that comes with humor, to mis portray things so that they will get really good at detecting falseness. Right. And, I mean, I know it works. And they know I'm not doing it for my benefit and against theirs. And so you know, it's rather like the example of your spouse having license. And in this case, it's even tighter because you know, you can lose a spouse, but your kids or your kids forever. So the question then is, what is the basis on which departures? You know, in effect, the Honorable shortcutting, or white lies that exists in the context of family are effectively a heuristic, right for family well being, in the same way that your perceptual apparatus makes jumps in order for you to process things in real time. So the question is, is there a basis of trust, and this is a long way of saying, the thing we can very clearly see when a for profit corporation is engaged in finding your perceptual defects and exploiting them is that a transfer of wealth is directly possible, right? You have the ability to decide where to spend your money, and where to spend your time, and they have an interest that departs from yours. And so they can effectively take time or direct wealth or whatever it is, in some direction, that you wouldn't have it go to your detriment, and they can come out ahead by doing it. And I think, you know, the, the ultimate upshot of this is if they do that, that it actually drains us collectively, something they didn't anticipate, right, the the, the side effect is potentially fatal to humanity. Correct.

**Tristan Harris** 37:35

Because I think we often focus too much in my work, especially when people reference time well spent started in this way. And people thought the only critique here was addiction or the harm was depression or isolation or one of these individual harms. But really the the emphasis on the in the long run is this sort of collective arrangement, which is, if you think about this is a business model that privately profits, but then collectively harms meaning it puts harms on the balance sheets of society and pollutes their balance sheet, you know, on the balance sheet of society is a inability to see a shared reality to have a different notion of history of what's happened, or what's fair, what's equal, because we've each seen different levels, literally infinite evidence of the other side lying or being hypocritical, and it creates no room for shared growth I wanted to bring back I know that your previous metaphor that I've loved about this is the notion of metaphorical truth versus the real truth. And the porcupine story, the idea that porcupines shoot their quills, I just thought this was so important in bringing it out, because this is a deception, but a deception that is in the greater benefit. Now the question is, who do you trust to construct that because you could have tribes that let's say, have false wisdom that are creating false deceptions? false stories that are shortcuts and heuristics. And then there's also a developmental aspect here, where maybe when you're young, they can say pork, you know, porcupine and shoot their quills. And so please stay away from Porky kinds. But when you're older, you've outgrown that, that metaphorical fiction and maybe I need a new one. And I'll tell you the truth, right? So maybe, you know, when you're young, I'll tell you about Santa Claus, when you're older, I'll tell you there's, there's a developmental new place to land. These are all interesting philosophical, I think areas of asymmetric persuasive relationships where we have, you know, more and more ability to influence people. But I also think I want to say it's not just psychological or perceptual vulnerabilities we're talking about here. There's also identity level vulnerabilities. And that's where this gets really insane. I mean, in the film, the social dilemma we talked about, how if I'm, you know, a foreign adversary of the United States, and I want people to not think this is not some partisan just Russia thing, or China is lots of countries that are now doing this. I can go into a conspiracy theory group, like acumen on like a pizza gate, or whatever. And I can actually get the user IDs of people that are that are of that mindset. Then I can pump them into a Facebook look alike model and say, give me users who have the same kind of profile like that. So Facebook is giving me it's not just that Facebook was doing gaina function research on mimetics or on groups or extremism. They actually offering keto function, it's been a function as a service, they're offering that to other parties to actually get more and more clever at being able to find and navigate. And then and then further their split test which messages would work for those audiences. And one of the things that we know, and again, I want to be really careful because we can steal man. And there's a pre trans fallacy on what a conspiracy is. And there's the sort of, you know, using it to deride people who think versus the actual genuine real conspiracies that are that are that are true. But I think this is really important because one of the other things we know about conspiracy thinking is the best predictor of whether you'll believe in another conspiracy is whether you already believe in one because it starts to work, the basis of your perception, that's kind of trust that you have, like, if it was true that the entire government would was lying about COVID. And it actually did come from the US. It I mean, which, who knows, who knows where it came from, it's topic of conversation. But if it were intentional in that way, that would certainly work the perceptions downstream of everything else. Because if that were true, then all these other things were true. Now we're getting we're getting into I don't want to go too deep into, you know, one of these specific ones or not, because I want to stay on the kind of identity hacking, but I wanted to make sure we were clear that the level of psychological and evolutionary hacking is mimetic, it's perceptual, it is social validation based, it is social relationship based. It's also identity based. In fact, in Facebook's own documents in 2000, I think nine or 10, that it was they had a Facebook marketing team in Australia. That said, we literally know when a child when when a user has low self esteem, they're talking about teenagers, they literally can predict based on usage patterns, when you are low self esteem. And that's an identity level manipulation, which is very much like cults, cults actually catch people in moments of vulnerability. What we know from cult recruiting is that they're actually in between huge moments of transition something in their life, a meaning making structure has collapsed, whether it's a relationship or a job, or a deceased loved one, or divorce. And they're in that moment of low self esteem and kind of low meaning. And that's exactly what makes them more vulnerable. And we've enabled and Sam Harris and I talked about this on his podcast, is a kind of cult factory, where there is a mass finding of people in vulnerable states, and then being able to personally target to a whole new degree, again, with a sort of gain of function metaphor, what what we could do to steer them. And the idea that this is the default infrastructure that runs our digital economy that this would allow, that we would allow for, and even make money from the gain of function zation of everything is incredibly dangerous. And we I just want to drive people there, because I want people to see just how deeply the reform is needs to be.

**Bret** 42:40

Yeah, marvelous. So I do want to briefly touch on this issue of conspiracy predicting whether or not you know, the belief in a single conspiracy results in a change in your likelihood of, or at least it predicts that you will believe in a second one, that actually this is about a heuristic being built. Right, to the extent that something persuades you that collusion has taken place in a given circumstance, it naturally adjusts your sense of how common collusion is, and corrosively, it adjusts your sense of how to weight evidence, because if there's one thing that's common to all conspiracies, it is that they portray some other possibility as more likely they inherent by frame some alternative. So basically, what it does is it forces you to sideline your obligation to outcomes razor, and there are cases where a little bit of sidelining of outcomes razor is absolutely necessary to make progress. But it's also the kind of thing that if you give up on it, you just become a sucker, right? That's right. So it's very hard to navigate that space. And basically, by pumping noise in, we are breaking people's relationship to reality and evidence very directly. And the consequences is all but inevitable, I would say, right, yeah, of course, people will just become so distrustful that they will effectively become full time superstitious.

**Tristan Harris** 44:05

That's right. And that's kind of where we are is what it feels like right? I mean, in the challenges you said is there there are ways of navigating the beyond Occam's razor explanation for something. And you're doing that a lot here on your program. And I think that's, you know, really fantastic but people need the tools to do that carefully because because there that process can externalize a kind of naive distrust or kind of naive epistemic process that starts to as you said, kind of get further and further superstitious if you don't have the tools and I think much like I always use a developmental paradigm when I think about these things, you know, you it's certain ages we speak in some ways it's other ages. We speak in other ways we tell you about Santa Clauses in our friend Daniel parquets point that, you know, he said life is a lifelong process of killing Santa Claus. Like you're sort of taught one Santa Claus metaphorical story, the porcupine thing, the simplistic story that suits you for a period it's a growth you know, it's a it's a jacket you can wear for that period, but then you outgrow it and there's a difference Mental paradigm. So maybe if we're going to engage in non Occam's razor narratives and really do investigations we have to be aware of developmentally which minds we're speaking to am I talking to someone who's literally kind of a basement conspiracy theorist with like a bunch of guns in their closet, and they're ready to go take arms to something consequentially, I have to be more worried about the way that my communication might impact that person, that if I'm talking to, you know, someone who's really thoughtfully engaging with alternative narratives, and epistemology, it actually has a whole process that they've developed to navigate that, and I think that's one of the things that we lose, especially in a social media landscape when you're broadcasting to, you know, 10 million people at once. We don't have a developmentally sensitive communication, meaning our communication is going to so many people that we assume, and this is probably the wise critique of engaging in any kind of alternate narrative thinking is, Hey, you know, you, Eric, and you, Brett are not sufficiently engaging with how people might hear you, consequentially. But then you're saying, Yeah, but you don't see that we're ignoring really important truths that we don't get to talk about, we're gonna get trapped by political correctness, or we're gonna get trapped by these other things. And this is where I think it gets very, very interesting.

**Bret** 46:07

Yeah, and actually, this is another uncomfortable topic, which is that I know, I am capable of calculating that some of the places where I question the official narrative, it is essentially certain that people are going to get hurt as a result of my revealing of these things are questionable. So for example, if we talk about what we don't know about vaccine safety, with respect to the the novel COVID vaccines, it is essentially inevitable given the size of the audience that a certain number of people will change their view on whether to get the vaccine, some of them will get COVID, some of them may die, they Other people may catch COVID from them who will die. That's a terrible responsibility. I don't want it no rational person could want it. But what I believe is that our sense making around these things is so broken, that we are already killing people as a result of our obligation to sign up for these official narratives that aren't true. And so I guess, you know, it's hard to know what the right analogy is. But the mechanism that you have been exploring, through which attention is captured, and sense making is deranged. That mechanism is like somebody has been working on a, you know, a computer of some sort of robotic device that you can put in the driver's seat of your car, and it can pilot you right? Now, I don't doubt that you can make a self driving car, we all know that's coming sooner or later. But if you tell me that you've built this in your garage, and you're going to put it in the driver's seat of the car, and we're going to go somewhere, we're going to crash, this is a prototype, right? And the nature of prototypes is that you're going to discover all of the hazards You Didn't Know About along the way. And you really want to discover that on a close track, you don't want to discover that with civilization in tow, right? And that's where we are is we are running the biggest experiment that has ever been run in real time. With all of us on board. There's not some population living somewhere who isn't downstream of the effects of this even if you're offline, right? We're all

**Tristan Harris** 48:17

totally it's, you know, you and others have said it's the largest unregulated psychology experiment we've ever run with no IRB board, if you want to use the Zuckerberg Institute of Neurology metaphor, and you have level what is it called the level five sort of IR app? Yeah, right level four. So if you have kind of the level four, you know, societal engineering lab, but you're literally not even doing it in the big white suits. And in the kind of enclosed spaces, you're just doing it on civilization constantly, you're in this sort of 24 seven, and we're actually still there, right? It's not as if right now there are not 3 billion people who just jacked into Facebook, Twitter, YouTube, etc, who are now in the live sort of, you know, mimetic tweaking of where these things are going. And so now the virus is sort of out of control. We're doing live tweaking to try to steer it, I think, at this point, by the way, I want to say I often am perceived as sort of a, I don't know a critic of the the system or making people evil, I don't believe any of these people are evil, I think they have not seen the consequences of the process that they were in the race to the bottom of the brainstem evolutionary landscape that they were competing within. But the problem is now they don't have you know, even the tools, the levers to actually do the gain of function, or that sort of the tweaking back towards some kind of stabilizing force. And that's I think the predicament that we're really in and frankly, this is the open inquiry I'm in is I think we obviously need tweaking of the technology and to fix it at that level because that's the thing that we're continually jacked into. But I also think we need almost like you know, once the pandemic of COVID is out there we you know, we do need some kind of protection mechanism to to, you know, isolate the virus, I think we need a cultural antibody, we need a cultural awareness. We need a cultural I don't want to say vaccine because that itself is polarizing, but we need some kind of common understanding and the thing that I don't want to speak too highly of the film's impact for the for the social limit. But I think one thing it does do, since about 100 million people haven't seen it is it creates a new shared common ground about at least one very primary explanation about how we lost common ground to this degree. So to have shared common ground about where and how we've lost Common Ground is at least a place to stand from, so that we can then say, how do we want to proceed back into more common ground? And I think that's kind of a, you know, an open inquiry that I'm in is what is what does that look like?

**Bret** 50:28

Alright, two things. One we are in if we go back to the earlier part of our conversation about shared fate, and what it allows you to assume is true. The irony here is that we have shared fate, right? This is a tiny planet, our industrial processes are at a scale now that we are our own biggest hazard. And so our fate is tied. And to the extent that our sense making is deranged, because in the moment, one group can get financially ahead of another by doing things that deranged us, in the long run, we may all perish together as a result of the process having not realized having not had it built into us that our fate is shared. And we should treat it carefully the way you would treat your spouse, for example. But I also think there's a way in which even the dire picture that you just painted, doesn't fully describe the hazard, because let's say that there was one, let's say that there was only Google. Right? If only Google was involved in deranging us, then Google stands a chance of figuring out the hazard and correcting things, right, knocking us away from the precipice. But the problem is, let's say that Google figured out that they were a hazard and realized they repented, right? And they, you know, invested a huge amount and figuring out how to under range us. But Facebook, and Twitter and Instagram, and all the other forces weren't on board with the program, right? It's a little like imagining, we're hurtling through an asteroid field, right? And you've got six people in six different control rooms, trying to adjust the cameras that allow us to see where the asteroids are. And so as one person, adjust a dial that makes the image better, somebody else may adjust the dial to put it back the way it was, because they were trying some other correction. Right? All of these disconnected things are basically an amplifier for noise. Right at its best. Yeah, at best. So now the question, really? Let's say that something like that is going on? The question really is what conceivable solutions exist here? Do you know?

**Tristan Harris** 52:47

I mean, this is our daily work, the Center for humane technology, and just on a daily basis, just among peers. I mean, I feel that the first point you're making about shared fate is exactly where we've just been the whole time. It's just we were all in this thing together, right? Because I think, to the earlier point we were making about we're all addicted, our amygdala are all available for mining and hacking, whether we know it or not, even if you're aware of the amygdala mining, it still works. If you show me something that makes me angry, I, I will get angry, I can try to be mindful of it. But again, the asymmetry of influence and the infinite feeds of the outrage are more powerful than my sort of daily mindfulness practice. And we shouldn't put that burden on individual minds anyway, to have to develop the superstructure or I think your brother calls it the metacognitive perch upon which to not be so hijacked by whether it's narrative control, or amygdala control. Um, yeah, I mean, when it comes to, I think we have to see some kind of shared fate. I mean, when you when you ask it that way, something I've been pondering recently, and you know, this history better than I do, because you were around when this happened, and I was just born a year later, is in 1983, the film The day after, comes to mind. I became obsessed when I my whole life with sort of the history of nuclear weapons and how close we got and how in the world after we invented the technology to destroy ourselves. I mean, there's all these nuclear scientists who were convinced seeing this exponential power, that they actually committed suicide. I mean, there's, I think the story of Is it fine Minar, one of the scientists who was in the back of a taxi looking at, you know, the bridges in New York City and just thought like, what's the point? Why are they building these bridges when it's, we already screwed it up, right? So it's over, we built the nuclear weapons, it's, you know, and they actually committed suicide thinking that was the end of the whole thing. And I think, you know, it's kind of remarkable because if you lived in that period, you probably would think, as I felt and if I was there, that it was over. And the interesting thing for those who don't know, this film The day after, was the largest primetime TV event in American history. I think it was a Tuesday or Wednesday night at 7pm on television. It was a film made by the director I forgot his name actually. But we'll look it up later. And it's about a hypothetical nuclear exchange between the US and Russia a full scale nuclear exchange I'm told through you know, narrative of regular people living in Kansas and a couple of Montana and a couple other places where the, you know, the missiles are underground, etc. And, you know, airing this on primetime television and having essentially the whole psyche of the entire country and it was aired around the world, essentially see the shared fate because we can't you know, if the share fade happens, it's too late. So we have to actually simulate the shared fate put it into the collective psyches of everyone so that the Omni lose lose outcome becomes visible to every actor, realizing that the escalation doesn't work. And this film was later credited with shifting Ronald Reagan's personal psychology because he saw the filament sent him chills gave him chills. It led to him wanting to do the Star Wars project, which was a failed project, but at least it led him away from the idea that we could win a nuclear war. And what's interesting about this film is it was aired on primetime TV 100 million Americans saw it largest TV event in history in the US. Then they had a panel afterwards under if you remember this, but they actually had a panel of Henry Kissinger, William F. Buckley, Jr. Carl Sagan, Brent Scowcroft, you know, the sort of the best military minds of the generation having a public discussion with the American public reckoning with Hey, hey, guys, we actually have had the power to destroy ourselves, we all just saw it. Now we're actually having the uncomfortable conversation. We can't be in denial about it anymore. What are we gonna do about this? And, you know, I just think that kind of film, that kind of initiative that kind of shared fate, visualization, stuffing that into the minds of everyone is the kind of thing that we need to ground, the much more extreme action we need to take that this will not work unless we take some action together. This is a shared fate.

**Bret** 56:57

That's beautiful. But there's a problem with it. I mean, I certainly remember the day after and, you know, basically, someone decided to traumatize us all together. I assume their motives were good. But the effect I think was certainly positive. as you point out, the problem is that in the case of nuclear weapons, and an all out exchange, the likelihood of coming out ahead, is pretty small. Right? Right. And so it is a much easier problem to solve by revealing its nature, because in general, the mindset that you have, you know, you basically get into Dr. Strangelove territory, where, you know, the crazies are interested in, you know, winning the post Apocalypse, right? So very few of us are on board with that. And so it's much easier to say, hey, actually, this nuclear weapon thing is a special kind of predicament that requires a new kind of cognition in order to even understand how to calibrate and of course, there's a lot to be said here. One thing that's true, is it's in some ways an unsolvable problem, until you have global agreement to get rid of the weapons because sooner or later, they will be used by accident or some other mechanism. So it's, we have temporarily solved the problem on a timescale that many of those who saw the invention of these weapons did not expect, it is also quite possible, and in fact, likely that the weapons themselves prevented a lot of other warfare, correct? Yep. Right. And so anyway, it's a kind

**Tristan Harris** 58:48

of accounting of the effects of these things. Right? Completely right there with you.

**Bret** 58:52

Right. But the other thing I would say is that the difference between the nuclear weapon situation and the derangement of civilization via food and search algorithms, etc, is it goes back to the question of magic. The fact that you know, that the magician is fooling you does not allow you to see in what way you are being fooled. And this is, you know, the The problem is there's no winner in the nuclear exchange, there is very definitely a winner short term in the the derangement problem. And so, we need some sort of, I mean, I think I think the point is, in some ways, you know, the social dilemma is the day after for derangement. And yet, as much as a lot of us sought understood the message agree on it. I don't know how much further ahead we are on the puzzle

**Tristan Harris** 59:52

completely. Well, I think this is this is critical, but I'm so glad we're talking about it step by step because I, where I'm lost, frankly, where I don't have an answer. As the outcome of the day after was he think about it from a theory of change influenced one person, I mean a set of leaders, especially Ronald Reagan, to shift policy, and then there was a shift in what could happen because those were people in control of a centralized exponential technology that could destroy everything into the shared fate. In this case, I guess you could say we have a small set of technology leaders. So those could be the, you know, the five Ronald Reagan's governing the five nuclear armed intersubjective, I think of the narrative derangement syndrome, you know, technology infrastructure we have as kind of nuclear weapons for the inner subjective world. So nuclear weapons were for the outer environment, but Facebook and Twitter and YouTube etc. Again, I want to, I also would like to make sure that your audience hears me acknowledge that I'm very aware of the incredible benefits that these things provide to and I just want to say, blood donors, finding each other high school sweethearts, finding each other YouTube producing lots of learning opportunities, music videos, pumping people up with exciting, you know, courageous, you know, ways of waking up in the morning, I see all that the challenges on the inner subjective side, going nuclear on a shared reality is a unsurvivable outcome if you can't recover from it. So there's two kinds of harm there's reversible harm where you cause the harm caused a you know, maybe a disease in the body. And there's, there's some kind of medical way to reverse it, or to rebalance your system symbiotically by, you know, whenever that's great. But then the irreversible harm the kind of I think losing trust, for example, in sense making losing trust, overall, not trusting any institution, not trusting each other, is a very hard place to recover from, that's in the more irreversible harm category. So what I worry about is that when you when you go nuke the inner subjective realm, especially at the trust layer, not just the inner informational layer. That's that's a place that I don't know how you recover from now, again, to your point, who would you convince so we have these kind of five ronald reagan type figures who are in control of the tech platforms, but they're governed, not just Reagan could theoretically make the decision himself. And yes, he had a country to report to, but he was no influenced by that country. In this case, we have five ronald reagan tech platform CEOs, I'm just throwing the number five around, there's lots of different tech platforms, but let's just use that. They report to their shareholders. They're publicly traded companies that actually can't get off this boat. So they're trapped just like we are, you know, and your brother, Eric has this great line that, you know, listening to tech company CEOs talk is sometimes like watching a hostage and a hostage video, the University a hostage hostage video, and it's like, they're saying things that don't really add up, you're like, why are they saying this gibberish, that doesn't make any sense. Until you see offstage, there's someone holding a gun to their head, and the gun to their head is their business model. And it's their, their share price that has to keep going up. And so that's, I think, you know, the kind of situation that we're in that we're in there.

**Bret** 1:02:58

Okay, two things. One, I want to put an asterisk around this question of shareholders, because I'm less and less convinced that it's playing the role we think it's playing. I know that it should, and I don't know why it's not. But I think we can see evidence that something else is afoot, and it may be somebody offstage with a gun to the head, but I don't think it's the shareholders at this point. But let's talk about the question of your analogy to a disease and the ability to restore the body to help. Yeah, the problem is analogous almost to half a bone marrow transplant, right? So you've got an immune system that's out of control, you've got leukemia or something, right. And the way to solve it, is to get rid of the immune system, which would be fatal in and of itself. If you didn't replace it with another one, you got to replace it, and it has to take, right. But if you just go halfway, the point is okay, you've just killed the patient, they were going to die of leukemia, and now you've killed them on a much shorter timescale. And in effect, the problem here is that the destruction of here, we've got a problem, you know, there's this this old, ironic line, you know, who you're going to believe me or your own eyes. What we've got is, you know, the irony of who you're going to believe me or your own eyes is that obviously, the answer is I'm going to believe my own eyes. Right? Okay. But the problem is, if somebody is gaming, your visual apparatus such that your eyes are not reporting accurately then the answer is actually, you know, I'm going to believe the mission the magician when they tell me what they did, rather than what I saw, right? In that case, I have to make an exception. Here, what we've got is a cyborg that is a fusion between intelligent creative human beings and unintelligent AI like stuff, learning to disrupt our own ability to manage our time, and to understand how to evaluate claims that we are seeing. And ultimately a smart person who is not in a position to control this phenomenon. And Nope, none of us are ends up distrustful of their own eyes and everything else. And the point is, you cannot just go back. Because once you've learned that your own eyes are lying to you, right? That your own instincts about what's true and false are unreliable. How does something use those instincts to inform you Okay, we've now restored sense making or sense making is getting better, right? The point is, once you've turned off the input, because you just don't trust it, you are now that much harder to cure. And it's very much like having had your bone marrow nuked with radiation, and then nobody bothered to infuse you with a new immune system. So anyway, I'm concerned that this is that there's an asymmetry here, where the damage is all too easy to do with this level of technology. And the process of undoing it is, you know, 100 times harder.

**Tristan Harris** 1:06:11

By the way, this is this is actually why there's no kind of kind of told you so phenomenon behind the words I'm saying now, but there's a, you know, so much what was so hard looking at this three, three years ago, I remember walking into senators offices talking about this, and just feeling like, Oh my god, there's no adults in the room. there's so few people who get that this process is happening. The more people who fall into these narrower, you know, more extreme views of reality is a hard process to reverse once you go there. And we have self reinforcing feedback loops, the gain of function process that is continuing to make this happen. This is going to be so bad in three or four years. And here we are. And I see that again, not there's no, you know, it, this is just why it was so important to try to stop it earlier. Because once you get here, it's very, very hard. Now there are epistemic processes to, you know, do more grounded sense making you can recover. I mean, I might one of my favorite teachers that Byron Katie has is sort of a semi New Age spiritual teacher, but she has a wonderful simple process you can use to just question your beliefs, it comes down to four questions for anything you're holding in your mind that that gives you stress, you just ask it's very similar to cognitive behavioral therapy, you ask, Is it true? Can I be absolutely sure that is true, so you're destabilizing the full grip of whatever belief that it is because no belief can fully represent 100% of the time, reality? or very few can? Then the third question is, how do I react? What happens when I believe this thought when I believe this belief? And then the fourth question is, who would I be? Not? What would happen like behaviorally? But who would I be identity level question without this thought, and you can start to kind of do culty programming on yourself, but deprogramming yourself from the cult of your own lifelong process of tribalism, identity, experiences, etc. and we can all be part of a culture that was trying to deprogram ourselves from the shared cult, or, you know, fractal cults that we've all been part of in this cult factory process. But as you said, We don't really have a lot of time here. And so that's where I just jumped to Okay, what can we agree on that are our problems that we have limited time on, and we try to figure out what will actually be sufficient there. And I know, you know, people like our mutual friend, Daniel Martin burger are people I trust to be doing that kind of process to be thinking about these various existential threats and societal problems, and what, you know, what, how could we get agreement on these things more quickly? It's very hard, you know, this is this is you could sort of, say Game Over, where where we are right now. But I, I prefer to say instead, what would it take? What would be most helpful? To align ourselves back on that track? And how can we align our actions on a daily basis, you know, towards that, that kind of world.

**Bret** 1:09:06

So, you know, I find myself in a very parallel situation. And in fact, it may be two facets of the same gym, right? So as you were trying to raise the alarm about the derangement of civilization and finding no adults in the room, as you know, I was watching my college go insane, and recognizing that it was actually a manifestation of a thought pattern that was sure to spill out into the world and start arranging us at a much larger level. And of course, the two things certainly interact, right. And there is this very frustrating feeling of, you know, I tried never to say I told you so. Right. It's hollow and pointless. On the other hand, I think what I want people to get from someone in your position or my position is Look, I clearly didn't know what was coming. Right? That could be a fluke. But it probably isn't because you can look at what I said about why I thought it was coming, right. And you can see that the thought process was right. In light of that, the fact that I'm currently alarmed about where we're headed, that I believe that the problem is dire, that as hard as it is to solve today, it only gets harder. Right? Those are things which should have are near complete attention because of what is at stake. And we do not need to know the details of what will go down, it doesn't matter. It's like not knowing how the robot that you've just put in the driver's seat of the car is going to crash it. Right? But you're gonna find out. And so anyway, you know, I have little things that I say to myself, and others if they want to hear them. One is, it's very late, but it's not too late. The point is, any part of you that says, Oh, my God, how are we going to solve this? It's hopeless? The answer is, well, if you think it's hopeless, now, try ignoring it another two years. Right? So getting people motivated to confront this now. And really, you know, your your point about Ronald Reagan, it's a complex point, but let's just say superficially, the point was, there was a Ronald Reagan, and he could be persuaded by a narrative, and it could affect policy. And unfortunately, you know, five major tech platforms governed by people who, frankly don't have the right kind of expertise to even navigate the hazard, right? They know their business very well, but they don't know how the rest of civilization functions very well. That this is this is urgent. And what we really need is human discretion. Right? There has to be somebody with veto power over the tech platforms ability to try any new experiment in order to do what they see as their job, right? There's too much at stake no matter what else is true. They are putting us all in jeopardy. It's not there, right? And it's obviously a dangerous and foolish experiment.

**Tristan Harris** 1:12:11

You're you're hitting on such great points here, because then then we get to some of the areas that you're closer to, which is, what does it mean for tech platforms to be governing all of civilization and doing so not in an autocratic way? And I know you've been the victim of, you know, sort of autocratic shutdowns in your own right. And moreover, you said, you know, let's take Apple, for example, because I've actually used the metaphor that Apple is kind of like the central bank or federal reserve of the attention economy. And they actually have the ability to do quantitative easing, or to maybe this is not the right metaphor, because people are against quantitative easing, for reasons I understand, but they have the ability to sort of tweak the rules, incentive guardrails, IRB standards, if you will, of the attention economy, Apple can say, Hey, we actually just don't allow this kind of psychological experimentation, you can do sandboxing, you can do proofs about what you think is going to help or not, we're going to name Make a list of what we currently deem to be just unsafe levels of harm or arrangement that is occurring from certain platforms. Now again, what I'm describing, you could say is what led to Apple taking down parlor and doing that autocratically and on their terms of, hey, you don't have enough content moderators? Well, ironically, due to the cultural derangement syndrome, people saw that in multiple different ways, depending on which part of the reality derangement, you know, fractal you were you were inside of so if you were on the right, you saw that as part of a left power grab and a left oriented, you know, political power within apple and Amazon, etc. being politically motivated to D platform, a right leaning technology product. If you were on the left, you saw it as something different. If you were trying to see it neutrally. And assuming that is not politically motivated, you could say that they actually didn't have the kind of moderation to deal with actual brewing problems that that they were not doing. So again, what would make apple trustworthy in setting up some kind of IRB process or in some kind of some kind of democratic Council for all of these issues. And as you said, the people who run Apple, you know, Tim Cook, Phil Schiller, Craig, I forgot his last name, who runs all of product at Apple. These are people with expertise in business with technology. These are not people who have expertise in civilization design, or essentially brain implants for 3 billion people, or a brain implant for the collective psyche of humanity, which is essentially the expertise they need to be in to make the decisions that are protective of humanity at this level. But I want to point to this because it's actually an area of success. Just recently, about a week ago, Tim Cook gave a speech. In fact, he said in that speech, we cannot allow a social dilemma to become a social catastrophe. And he was referencing the social dilemma in in reference to a set of changes that Apple was making to go after him. Targeting business models, the ability for apps to track you across different applications. In other words, sort of taking down the advertising business model in its micro targeted form. And moving back to kind of a 1960s 70s era, billboard advertising, contextual advertising, topical advertising, but not micro targeting. So this is Apple essentially saying, Hey, we have this whole ecosystem. And we're going to nudge the ecosystem away from some of the toxic components of the attention economy towards a less personalized, not micro targeted form of the attention economy. Now, this is a very small change. But it deeply is affected many of especially the kind of Facebook and Google parts of the world, which they're very, not happy about and claiming anti trust, sort of motivate, you know, actions against against apple and taking these unilateral action. But I think we can zoom out and say, okay, there's two ways to fail here. One is, you know, autocratic tech CEOs wake up and just make, you know, instant reactionary decisions based on the impulses of the present moment. And that's bad. But the other way to fail is to not make decisions about things that are actual threat. So we don't want to not make a decision and allow the Frankenstein to keep eating up the world and arranging the world. But we also don't want autocratic decisions where some algorithm gets produced. And suddenly, Bret Weinstein is off of Twitter and Facebook, and there's no appeal process and you happen to you know, hear the person from Facebook get back to you later, compared to the 1000s of people who would have no appeal process. So what we need is some kind of, you know, there needs, it needs to be in a governance of humanity, essentially some kind of democratic governance with the expertise of what it means to be a brain implant for the collective psyche of humanity. Apple's in a very unique position because it governs the App Store and Google because it governs the Play Store on Android to try to steer things in that direction. Some people are gonna like the sound of that, because they don't like the power that that that what that means. But I think we have to say it's, we have to have some governance process.

**Bret** 1:16:52

Yeah, it's both hopeful, because you point out that the power exists there and hopeless. I mean, even just the idea that we have to prevent this from becoming a, what did you say? derangement catastrophe? That idea is, I don't know, six, seven years too late. We're already there. Right? It's the choice between a crash and a crash landing. It's not a hard landing at this point, right? We've got a serious problem. And nobody knows how to turn the ship around. And so it's a little bit like, you know, the captain of the Titanic, beginning to grapple with the hazard of, you know, a iceberg as they show up on the horizon. It's like, well, it's the goddamn Titanic, right? It's not going to turn on a dime, you need to wake up faster, right? And there's no there's also the point that okay, Apple finds itself in the position that you describe. It doesn't deserve the position you describe, it doesn't have even the superficial characteristics that you would imagine would allow it to navigate that position. Well, in other words, okay, what is it that's supposed to tell to inform apple of the way in which it is supposed to steer things away from dangerous derangement phenomenon? Is it shareholders, they are obligated to maximize shareholder value, which I already told you I suspect is not playing the role, we think because we see corporations now routinely making decisions that are nakedly political about who they don't want on their platform, even though it's hard to imagine how that enhances their, their bottom line, right? You would expect a kind of indifference about content in order to get everybody on the platform and not create a market of a competitor for a competitor to, you know, leap up and serve customers who've been thrown off the platform, you would expect a lot of things that we're not seeing. So what is that? Is that shareholders deciding that they value something above financial return? Possibly, in which case this is novel territory? I would say, maybe I'm missing something. But the idea is, you know what? Apple can't do it. Your right. Global processes need the global governance and of course, this freaks people out. But my point is not we need global governance. My point is, global processes need global governance, right. The alternative to a governed a well governed global process is an ungoverned global process, and that is a disaster.

**Tristan Harris** 1:19:35

So it's where we are now right? And without taking any actions. We're in that dystopia. If we take actions. We're an autocratic, 1984, dystopias, we need something that is actual global governance. And, you know, and you're right to say, you know, it's not like we would have wanted to be in this situation where this particular set of human beings ends up at the top of this decision making hierarchy because they had the expertise No, they got there because they were good at building it. Grabbing tech products or the hardware software, you know, 40 year trajectory of Apple, but you know, how we how we navigate from there what that looks like. I mean, what, you know, I, Justin Rosenstein, who's in the film and has a new book out called the new possible I wrote an essay for it and it's it's actually about visions of the possible from where, you know across the sort of economic landscape where we see the systemic problems and one of the things that's brought up a lot is sort of the citizen council or citizen juries to actually, you know, the James Fishkin work on deliberative democracy, bring groups of citizens into a process where they actually hear all the evidence from experts, and they're actually, there's it, it's proven to lead to actually pretty even dramatic changes in people's understandings and beliefs in an actions about what they would want to happen. In Ireland, for example, there was a whole thing a referendum, I think, on gay marriage, and it led people who had even been against that in the beginning to, through, you know, a two or three day process lead to actually voting, I think for legalizing I think gay marriage. And we could have processes like that. But again, as you said, we have to sort of global, we have a global thing. So we do that in English, you know, how many courts do we need? How many juries do we need? You know, one of the things I know I was planning to be ready to talk about his, with your situation with getting d platformed. Without any kind of review. You know, we need appeals processes and courts. But if you think about the scale of accounts, the scale of messages that are needing to be adjudicated, we don't have juries and courts for enough of the people messages, etc, that are going to be flagged, we're generating more, you know, discrimination or cases than we have the ability to have courts to, to govern. And I think this is actually a fundamental paradox in technology, because the benefit of technology is automation at scale, the fact that you have fewer human being decision makers, and you get more leverage through these automated processes. But when a human judgment situation comes up, we need to have as many courtrooms and lawyers and judges and juries, as we have issues that we need to be adjudicated. And if we have things that are consequential enough to cause irreparable harm faster than the court case, date comes up, or your case can be heard. You know, you don't want to wait you know, a year two years, five years until you get to decide whether you get to come back to Facebook or Twitter. We don't want to have Apple wait years to make decisions about you know, the derangement syndromes and the IRB standards that they would try to put to you know, enclose these these perverse Frankenstein's I think we need a rank priority list though, of how we deal with these issues and protect against all the forms of tyranny and dystopia that we want and you know, you've many times pointed out the sort of Orwellian tyranny that you're that you were the victim of. And then the other tyranny we've been talking about for much of this conversation so far has been the Huxley in tyranny, Huxley and dystopia, where it is not in this is stolen from Neil postman's, amusing ourselves to death. But we always worried about 1984, and the ability to ban books and censor information and censorship. But then we didn't worry about this other world where we give people so much information that no one actually reads books, no one actually has the attention spans, everyone is sort of stuck in their pleasure machine in their affirmation machine, there's so much machine that they don't have a legitimate society. And so I think if we make some lists of here's the dystopia that we're either in or we don't want to fall into here's the ranking of those problems. What are the things that need to happen at a global governance level to do it? You know, this does look pretty bleak. And let's be honest about where we are. You know, I, you and I have talked in the past about how hard it is, psychologically to wrestle with staring these things face to face on a daily basis. And I think that's actually part of this is psychologically How can you look at some of these things, but I think that's where we are, we have to ask, what would help from that point, and I'm curious if you see things differently, or

**Bret** 1:23:44

No, I mean, I think you know, you and I have a difference in the way we look at this. I think we both see a bleak picture. And I think there's an emotional difference in how we react to it. But there's not a substantive difference. I think the the number of dystopias that we can allow us, ourselves to fall into is zero, right? We agree that there are many and that are there for tensions between solutions. But I want to divide two things because you talked about the need for you know, an a, a vast court system in order to address all of the potential claims of an infraction, etc. That's one way to do it. The other way to do it is to separate the question of the derangement that arises through algorithmic competition. From the question of speech entirely, right? In other words, I don't know of an effort to not sell pens, ink, paper, any of these things to people with a political perspective, in other words, I think We have swallowed the idea that actually we're not going to prevent Nazis from rising by denying them paper. Right? It's not a good bet, right the world you have to build to keep Nazis from getting paper but allow everybody else to have it is not a tolerable world. So the point is okay, the ship has sailed, they're gonna have paper that's not we have to fight them on the landscape of ideas, not the landscape of writing implements. And the question is, what is the online environment? Well, it's two things, it's a bunch of businesses battling and putting us in the danger that you're talking about. And it is also the public square, whether you like it or not. And so at the level of the public square, I want to know very clearly what the limits are of what can be said. And I do not want people policed for even obnoxious points of view, right? obnoxious points of view are going to have to be tolerated. And we have to trust in the fact that given complete freedom to talk, the obnoxious points of view are driven out bigotry declines, right? That's been the nature of it, and trying to police bigotry is a recipe to create exactly the opposite phenomenon. So So anyway, let's separate that out. And let's say the way to deal with the unimaginably large number of courts that we would have to create to police every infraction is to say, actually, the infractions are going to be fairly clear and few in number, because what you're going to be allowed to say online is going to be the same thing you would be allowed to say, standing on a street corner, right? And then we can separate off the issue that's really a hazard to us, which is, so we create a hazard if we try to police speech online, right? There's no conceivable system that could do that. Well. And the dystopias that arise from bad attempts to do it are several. But then we've got the other dystopia, which is the evolutionary novelty dystopia, where something decides to borrow, you know, something decides to play magician, and it begins to evolve. And before you know it, nobody even knows how it works much less how to regulate it, there has to be some body of people. And here's the hard part to sell to people who have absorbed, the true lessons of the magic of the market is somehow the people with discretion to say, what we cannot do, what we must not be allowed to do to each other using technology have to be independent of this, you can't be, you know, a wealthy person with a stock portfolio in which you're invested in all of this stuff and decide whether or not it's smart to engage in these experiments, the perverse incentives are too many. And people are not even good people are very difficult at managing perverse incentives. So I don't know how we get there. But I do have the sense that in the end, either you have to forbid the market to govern anything, or you have to have somebody who's immune to market forces in a position to say, actually, you can't do that, because you're going to deranged people and you're going to result in a tremendous amount of harm, right?

**Tristan Harris** 1:28:09

You can't do what?

**Bret** 1:28:13

x? In other words, if somebody had proposed Facebook, and the idea is Oh, well, Facebook is a place where people can gather across arbitrary distances and talk about topics they want to talk about. Okay, so far, so good. And, you know, they'll be able to see a feed of what their friends are doing, saying, etc. Okay, so far, so good. And now that feed is going to be adjusted according to an algorithm that, and then the answer is, well, what is that? Right at the point you begin to propose the capture of attention, right? Something needs to say, wait a second, here's what happens downstream of that, or, you know, what? We don't know what happens downstream if that, right. So in other words, if we go back to your analogy about gaming, again, of function research, we had a running battle about whether or not gain of function research was necessary in order to protect us from a pandemic or was more likely to cause a pandemic than sorry, one, right, right. Those of us who were on the latter side of that didn't win. And those of us who were on the Hey, we got to do this I did. Now is that the reason that we are now fending off COVID-19 and losing the battle, frankly, somebody needed to say Actually, no, because not necessarily. Here's the dystopia that arises if you do this, but we don't know what arises if you do this, right, because you're a principal, yeah, precautionary principle, which is, I admit, difficult to instantiate well, but some version of it has to govern us.

**Tristan Harris** 1:29:52

And this is kind of an age old argument between you know, allowing, markets are just another form of evolution. So do you just allow system and technology To just, you know, let people build things and, you know, fix it later versus when do those things have the capacity to cause irreparable harm. And we should instill the precautionary principle and say, let's not just jump ahead and you get the kind of technical libertarians that just want to, you know, build, build, build no regulation, deregulate everything. But again, I think what we're finding is, we have to be able to spot when the things that we're proposing building are closer to this sort of, I don't keep referring to get a function, so people don't center on that at the center point of our whole conversation here, or what it could mean, but I just that there are dangerous forms of tinkering, that if you're doing social tinkering, or mind, you know, meaning making tinkering or social group, tribalist tinkering, which all of these things Facebook had engaged with, you could cause some, you know, some, some some real harm. And there has to be some process by which we stopped that. But there was another thing that you were saying about speech in the Nazis in the paper in the pens, I want to make sure we're at least acknowledging a principle that we can, I think, agree to and want to honor, which is that, you know, to whom much is given much as expected, or with great power comes great responsibility. And the greater the power, the greater the responsibility and accountability. And we went through a decoupling of that. And I'm really borrowing this from Daniel shmotkin burger, and this is this is his insight, not mine. But you know, I can go to a kitchen store as a Nazi or as a regular person, and I can buy a set of knives, knives can be used to kill people, but they don't do a background check. They don't force me to do knives training, they don't have, you know, I can just do that, because there's a limited degree of harm that I can instantiate as an individual. And even though it sucks, we have a court process to later be able to find those people, etc, whatever. But if I'm going to buy a gun, we have, you know, Second Amendment, so I guess we have a fair equal access to guns. But we also have background checks, we have ID checks, we have sometimes cases of you know, training, and that's required to wield these kinds of weapons. But if you're going to buy a tank, or buy a bio weapon manufacturing system, or ammonium nitrate, or what these other things that we don't just say, Hey, everyone gets access to that, or if you're going to, again, do gain of function research, or have access to ammonium nitrate, or have any of these other things, we have to make sure that we're putting protections around it. So now with speech, I think we have this temptation to say speeches, just speech, but if I'm actually broadcasting to 100 million people, and I'm ability to do that unregular with no regulation No, no limits, no. You know, I can say whatever I want, I can say it with the most malicious intent, I can, I can do consequential analysis, I can run, you know, gainer function, you know, simulations on what I could say that would cause different tribes to react in the exact same way I can literally build a machine that's generating speech and split testing it across trillions of simulations, knowing the communities that respond to my messages, in which ways and I can engineer almost like lab generated red meat, political red meat, that I'm going to throw into the ecosystem. And so I'm in other words, I'm doing kind of gain of function, political research on what would cause the most havoc, the fact that I can do that degree of genetic manipulation, and caused global chaos at that level. And there's literally no distinction between the responsibility of someone with 100 million followers on Twitter and someone with 10, I think we've lost the principle of to whom much is given much as expected, or at the very least, if you compare standards for broadcast abroad, TV broadcasters who reach maybe a million people in a local area, they face more regulations and standards, then we grant online. And so I think these are new kinds of services. Even Facebook is very different than Twitter in terms of text communication, versus, you know, the kinds of posts that you can make on Facebook. But we're certainly missing that principle that the responsibility accountability standards and practices are needing. And I would put that in the list of the agenda of the constitutional convention that I believe is necessary to make this civilizational infrastructure, more democratic and globally governed for the public interest and public good.

**Bret** 1:34:01

Well, I agree with a lot of that I sort of feel like we need a constitutional convention, and we certainly cannot afford to have one, it will end up a disaster if we do so the answer is we need something to reckon with things of that scale. Yeah. One of the things that I say about the the market, what what is the market? What can we allow it to do? And what Shouldn't we is that there's a natural distinction between the following things. Markets are excellent at figuring out how to do things, and they are appalling at figuring out what should be done, right? But they tend to do things like exploit defects in your personality, find weaknesses, amplify insecurities, in order to make money, and we can't let them do that. It's not it's not kind it's not decent. So, right, so there has to be some sort of human discretion on what do we allow the market to do. And then there has to be liberty to allow the markets to do things, but the other thing that I'm increasing compelled of is that we have a process in which once things show themselves to be profitable, they cannot be reversed, right that the political dynamics are such that once you've made a huge profit at something, you're unstoppable, so we can't go in reverse. And that is in and of itself an existential threat. I can, I won't, I won't do it here because we're pressed for time. But that will actually cause the equivalent of civilization senescence, it will grow feeble and collapse because of that dynamic. And so, I mean, it is amazing. Look at the technology we are using to have this conversation, right, we are almost as if face to face having a conversation in real time, which will be broadcast to hundreds of 1000s of people. Okay, that's an amazing level of technology. And yet, we cannot figure out how to stop a technological process that everybody smart understands is putting us at existential risk. You can unplug it, right. And you know, the problem is even to you and me, I'm sure the idea of tech platforms being unplugged, because civilization depends on them being unplugged is almost unthinkable. But how did we get to the place where there are processes that we have unleashed in the world that are both a threat to us, and inconceivable that they could be unplugged?

**Tristan Harris** 1:36:21

Yeah. There are days when I just think we need to unplug these things. And I actually want to make a distinction, because what we're doing right now is a good example of, I think, a much more humane technology in the sense that our representation to each other, the fact that I'm making almost eye contact with you the fact that my trustworthiness and the credibility of what I'm saying and how I'm saying it would be very different than I think when you think about the medium of Twitter or Facebook, if you were sort of harboring really crazy views or whatever, you would have almost an inflated sense of credibility talking about those things in those medium because they subtract away, how credible does that person look when they're saying all those things. And I think if you were to reverse the last 15 years, and you just had zoom conversations, you had zoom conversations that were deep like this, that are two hours long, you had even 1020 people on them? Would we see the kind of mass bifurcation of reality? I mean, we still would have partisan media, I want to make sure acknowledging all of that, etc. but would we have the mass, you know, fractal, you know, fracturing of reality? Would we have the degree of sort of out there crazytown kind of thinking that we have, would we have, you know, I just don't think we would have those things, because we would be in these long form mediums where we can say, Wait, hold on, but I want to make sure I understand you correctly. Were you saying this, are you saying this? Now we need modeling of that kind of turn taking too, we need modeling of good epistemic process. But that's, I think the thing that you're doing with this podcast and that others are doing and I think this kind of medium is much more humane than something that is an automated gain of function research, sort of, you know, plant tinkering with human evolutionary instincts, and then letting things go, you know, basically pandemic human medics. It's a pandemic mimetic complex, that's what it does is it creates this sort of global spread of information algorithmically. That was not safe this summer, this is much more akin to safe. And I would prefer if I could go pull the plug, and you asked me, and another thing we saw each other, I think about two years ago, you know, you'd be one of the few people I would trust to kind of pull the plug in, in some ways, I think this would be a plug I would be willing to pull is to take away this sort of algorithmic social media, and leave ourselves with this sort of long form, you know, podcast, like long zoom conversation, like with even 100 200 people listening and interacting, that would be a much better world to live in.

**Bret** 1:38:36

This is gonna be a marvelous point to end on, because it sort of wraps up much of what we have been discussing, because the point I would add to what you just said, with which I 100% agree, is that there is a reason this is safer. It's not arbitrarily safer. It's safer, because this treads, it is fully technological in every way. But it allows your built in heuristics as a human being to function. Right? So in other words, somebody who was standing three feet from you and me talking at a cocktail party, were we to have the same conversation would have the same ability to evaluate. Are these people full of shit? Did they know what they're saying to each other? Are they attempting to persuade others of their credibility? Do they have a financial interest here that isn't obvious, you could make that evaluation in a normal room. And it would, you know, wouldn't necessarily be right but it would be plausible that you could do it you have enough information to go on. And so anybody tuning into this conversation will be able to detect, for example, they may not perfectly be able to detect whether there are any edits. I don't think they're going to be any edits here but they will be able to see whether the conversation flows as if on edited. They will be able to tell that you and I are responsive to each other that neither of us is working from talking points because each point follows from the next in a very natural way. And so the point is your your built in set of heuristics for how human being sound when they do or don't know what they're talking about is applicable here in a way that it can't be when things are spliced together. And an algorithm is playing an unknown role. And the algorithm is changing. And you can't even detect that if your heuristic worked yesterday, it might not work tomorrow, right? All of those things are the challenge. And the closer things look just as with diet, the closer things look to what your ancestors have been eating, the more likely they are to be well dealt with by your system.

**Tristan Harris** 1:40:36

And by the way, everything you're describing when when we use the word humane in our founding of the Center for humane technology, it actually comes from my co founder, Eva Raskin, his father, Jef, Raskin started the Macintosh project at Apple. And the primary notion that humane means being respectful of human vulnerabilities and essentially understanding what was what was adaptive about all the evolutionary architecture that you would want to express in its fullest capacity, it helps in a healthy way, that that's where we need to look to get this kind of dimensionality this kind of that outcome, like you said, so just to say that with the philosophy you're talking about is exactly what we say when we think about you have to be looking not at what would make technology more sophisticated? How can we be more sophisticated about human nature? And the evolutionary roots that got us here in a healthy way? And how can we reappropriate more and more of those lessons, I think that there's interesting people across the board that are doing this, I think we need to do this kind of mass excavation for the babies, we threw out with the bathwater of other cultural traditions. I mean, I think even some of Jordan Peterson stuff of taking personal responsibility, as simple as it is, these are kind of wisdom hacks that bolt on to the human evolutionary system. One more that my might leave you with that I one of my favorite ones recently, is the Serenity Prayer. The notion that you know, God, give me the wisdom to, I'm going to botch it, what is it God give me the wisdom to know the things that I can't change to accept the things that I can't, and have the wisdom to tell the difference? If you think about that, I'm curious if you agree with from an evolutionary perspective, in the evolutionary ancestral environment on the savanna, attention was coupled with agency, I put my attention on that rock, I can move that rock my attention, I put that on that lion, I can run away from the lion. So what am I put my attention on is directly coupled with what how I can act in relation to it. And so our brains don't naturally distinguish, we put our attention on things between what we have agency to do something about, versus that, which we don't. And that's why, you know, I think of that, that Serenity Prayer, that wisdom prayer as a sort of bolt on humane technology that says, hey, your mind in this new environment, doesn't make that distinction. So we have to remind ourselves to point our attention to things that we can change, and humane technologies like that it's based on an insight about the fact that our attention is agency blind or non discriminating. And that we need help to put our attention on the things that we have agency over just to leave the audience maybe with an optimistic vision is imagined a humane social media or, you know, other sort of tech digital technology that is helping us put our attention on the things that we can actually change, and that ladder up to bigger changes. In a world in which every single time we use technology, it's helping us put our attention on things that can actually change we can get on that treadmill of positive feedback, of feeling like the daily choices we make are leading to a better world for ourselves, for the people around us and for the world. And that kind of attention economy is a much better attention economy than the one we have today, which is basically infinite learned helplessness. Because even if I'm pointing my attention at the important stuff, like say, climate change in our Facebook newsfeed, it's basically Yeah, it's worse than you thought. And there's nothing you can do. And I can scroll my finger and see, it's worse than you thought nothing you can do worse than you thought at the you can do. We want a world where we have agency over the things we've put our attention over and technology could make that the default setting, as opposed to something you have to consciously bolt onto your own consciousness.

**Bret** 1:43:55

Wonderful, I want to just point out the flip side of it. So you're absolutely right, that our attention is essentially scheduled to those things that an ancestor would have had agency over and they would be programmed, they would be wired to ignore everything that there was no point in thinking about it, because there was nothing to be done right? Now the problem is, if you're Mark Zuckerberg, then you could make decisions about algorithms that could plausibly end the world. So your attention really be ought to be on that hazard in a way that an ancestor would never think about it because there's nothing an ancestor could do that would end the world. Right? Right. Right. There's no decision that works that way. Or very, very rarely, you know, you could maybe screw up your you could take your population across some something or other and go extinct, but in normal functioning, there was no mistake you could make that would have that kind of impact. And so I need to get these people upgraded. So they understand the increased agency that they have and therefore the increased responsibility which I realize is part and parcel of your your formulation here. So anyway, okay. I feel like we've had At the introduction to a very good conversation that we need to have,

**Tristan Harris** 1:45:04

we should definitely do something like this again, some time. It's been a fantastic conversation, I really appreciate getting to dig into these topics with you and go deep.

**Bret** 1:45:12

Great. It's always great to talk to you about this. And just to say, in closing, I do think there are people who can be trusted to think this way on humanity's behalf. And you are very high on my list of that small fraction of people who has that kind of wisdom and decency. So anyway, thank you for all you do Tristan, really appreciate it. And I look forward to the next one.

**Tristan Harris** 1:45:38

needed lots of meat bread, thank you so much. Where should people find you? You can look up more of our stuff on humane tech comm find me on Twitter at Pinterest on Harris, although I don't really use it very much given kind of post social dilemma world. But I think our podcasts, your undivided attention is a great place to look for for going more deepness in these topics.

**Bret** 1:45:59

Great. That's excellent. And we will post links to those in the description and be well Tristan and to everyone else. Thanks for tuning in. Thanks so much.