

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

- Tristan Harris: Hi, everyone. It's Tristan, and this is Your Undivided Attention. Up next, we have our unedited conversation with Daniel Schmachtenberger, and because it's unedited, it's longer and not corrected for fact-checking purposes, but you can find our shorter, edited version wherever you found this one. Listen to both versions, and then come to our podcast club with Daniel and me and hopefully you on July 9th. Details are in the show notes, and with that, here we go.
- Tristan Harris: Welcome to Your Undivided Attention. Today, I am so honored and happy to have my friend Daniel Schmachtenberger as our guest who works on the topics of existential risk, and what are the underlying drivers of all of the major problems or many of the major problems that are really facing us today as a civilization, be it climate change, breakdown of truth, social media, or information systems? Those of you who've been following Your Undivided Attention will hear this as a very different kind of episode. We almost think of it as a meta episode about the underlying drivers of many of the topics that we have covered on Your Divided Attention thus far.
- Tristan Harris: So, if you think about the topics that we've covered, whether you've seen the social dilemma or you followed our interviews previously on topics like attention span shortening or addiction or information overwhelm and distraction, the fall of trust in society, more polarization, breakdown of truth, our inability to solve problems like climate change, well, this is really about an interconnected set of problems and the kind of core generator functions that are leading to all of these things to happen at once. So, I really encourage you to listen to this all the way through, and I think that we're going to get into some very deep and important knowledge that will hopefully be orienting for all of us.
- Tristan Harris: One of my favorite quotes is by Charles Kettering who said that, "A problem not fully understood is unsolvable, and a problem that is fully understood is half solved." And what I hope we talk about with Daniel is, what about the framework that we are using to address or try to meet the various problems that we have has been inadequate, and what is the problem solving framework that we're going to need to deal with the existential crises that face us? So, Daniel, welcome to Your Undivided Attention.
- Daniel Schmachtenberger: Thank you, Tristan. I've been looking forward to us dialoguing about these things publicly for a while.
- Tristan Harris: Well, you and me both, and for those who don't know, Daniel and I have been friends for a very long time, and his work has been highly influential to me and many people in my circles. So, Daniel, maybe we should just start with, what is the meta crisis, and why are these problems seemingly not getting solved, whether it's the SDGs, climate change or anything that we really care about right now?
- Daniel Schmachtenberger: I think a lot of people have the general sense that there is an increasing number of possibly catastrophic issues, and that as new categories of tech,

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

tech that allows major cyber attacks on infrastructure, tech that allows weaponized drone attacks on infrastructure, biotechnologies, artificial intelligence, and moving towards AGI, that there are new catastrophic risks with all of those categories of tech and that those tech are creating larger jumps in power faster than any types of jumps of tech, including the development of the nuclear bomb in the past by many orders of magnitude.

Daniel Schmachtenberger: So, there's a general sense that whether we're talking about future pandemic related issues or whether we're talking about climate change or climate change as a forcing function for human migration that then causes resource wars and political instability or the fragility of the highly interconnected globalized world where a problem in one part of the world can create supply chain issues that create problems all around the world, is that there's a sense that there's an increasing number of catastrophic risks, and that they are increasing faster than we are solving them, and that when you mention like with the UN, while progress has been made in certain defined areas of the Sustainable Development Goals and progress was made back when they were called the millennium development goals, we're very far from anything like a comprehensive solution to any of them. We're not even on track for something that is converging towards a comprehensive solution.

Daniel Schmachtenberger: And, if we look at the core initial mandate of the United Nations in terms of thinking about how to, recognizing after World War II that nation state government alone wouldn't prevent world war now that world war was no longer viable because the amount of technology we had made it a war that no one could win, we still haven't succeeded at nuclear disarmament. We had some very limited nuclear disarmament success while doing nuclear arms races at the same time. We went from two countries with nukes to more countries with better nukes and that simultaneous to that, every new type of tech that has emerged has created an arms race. We haven't been able to prevent any of those. And the major tragedy of the commons issues like climate change and over-fishing and dead zones in the oceans and microplastics in the oceans and biodiversity loss, we haven't been able to solve those either.

Daniel Schmachtenberger: So, rather than just think about this as like an overwhelming number of totally separate issues, the question of, why are the patterns of human behavior as we increase our total technological capacity, why are they increasing catastrophic risk, and why are we not solving them well? Are there underlying patterns that we could think of as, as you mentioned, generator functions of the catastrophic risk, generator functions of our inability to solve them, that if we were to identify those and work at that level, we could solve all of the expressions or symptoms, and if we don't work at that level, we might not be able to solve any of them?

Daniel Schmachtenberger: Again, people who've been thinking about this for a long time notice these issues. They notice that you try to solve a ... The first one I noticed when I was a kid was trying to solve an elephant poaching issue in one particular region of Africa that didn't address the poverty of the people that had no

**Center for Humane Technology | Your Undivided Attention Podcast**  
**Episode 36: UNEDITED A Problem Well-Stated is Half Solved**

mechanism other than black market poaching, didn't address people's mindset towards animals, didn't address a macro economy that created poverty at scale. So, when the laws were put in place and the fences were put in place to protect those elephants in that area better, the poachers moved to poaching other animals, particularly, in that situation, rhinos and gorillas that were both more endangered than the elephants had been. So, you moved a problem from one area to another and actually a more sensitive area.

Daniel Schmachtenberger: We see this with, well, can we solve hunger by bringing commercial agriculture to parts of the world that don't have it so that the people don't either not have food or we have to ship them food? But if it's commercial agriculture based on the kind of unsustainable, environmentally unsustainable agricultural processes that lead to huge amounts of nitrogen runoff going into river deltas that are causing dead zones in the ocean, that can actually collapse the biosphere's capacity to support life faster, then we're solving for a short term issue that's important and driving even worse long-term issues. We see that many of the reasons people who oppose climate change solutions in the West, oppose them is because not because they have even really deeply engaged in the underlying science and say the climate change isn't real. That will oftentimes be what's said, but because of the solution itself seems like it'll cause problems to other areas that they're paying attention to that seem even more critical to them.

Daniel Schmachtenberger: So, if the solution involves some kind of carbon tax or something that would decrease GDP for the countries that agree to it, and some other countries don't agree to it ... Let's say in this particular case, the model that many people have is Western countries agree to it. Their GDP growth decreases. China doesn't agree to it, and there's already a very, very close neck and neck fight for who controls power in the 21st century? Are we ceding the world to Chinese control that many people think it has less civil liberties and is more authoritarian in its nature? Or some people's answer to climate change is, well, we just have to use less energy, but when you understand the energy correlates directly to GDP and when GDP goes down, it affects poverty, people in extreme poverty first and worst, and wars increase because people who have desire to get more end up going zero-sum on each other, and only when it's very positive sum does that not happen. You see all these intricate theory of trade-offs, so we can't see that the problem is climate change.

Daniel Schmachtenberger: Everybody knows the problem of climate change is like a big thing, but you've got to look at climate change plus the macroeconomic issues that would affect the poorest people and that would increase the chance of war and the geopolitical dynamics between the West and China, whatever, and the enforcement dynamics of international agreement. When you start to recognize that the problem is that suite of things together, in a way, it seems, well, that's too hard. We can't even begin to focus on it.

Daniel Schmachtenberger: I would say that that's actually easier. Because trying to solve climate change on its own is actually impossible. Because if you're trying to solve

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

something that is going to externalize harm to some other thing, maybe the other thing that you ... Maybe you solve that thing, but you find out that you're in a worse position. So, I would say that it's impossible to actually improve the world that way. Or half the world that is paying attention to that other thing disagrees with you so vehemently that all the energy goes into infighting, and whatever some part of the world is trying to organize to do, the other part of the world is doing everything they can to resist from happening. Then, all the creative energy just burns up as heat, and we don't actually accomplish anything.

Daniel Schmachtenberger: So, I would say that the way we're trying to solve the problems is actually mostly impossible. It either solves it in a very narrow way while externalizing harm and causing worse problems, or makes it impossible to solve at all because it drives polarization. And so, going to the level at which the problems interconnect where that which everybody cares about is being factored and where you're not externalizing other problems, what seems more complex is actually possible, and possible is easier than impossible. So, it's not just that there's a lot of issues, right? There are a lot of issues and just that the issues are both more consequential at greater scope and moving faster than previous issues because of the nature of exponentiating technology. That's part of it. It's not just that the problems are all interconnected. It's also that they do have underlying drivers that have to be addressed. Otherwise, a symptomatic-only approach doesn't work.

Daniel Schmachtenberger: The first underlying driver that, when people look at it, they generally see is they see things like structural perverse incentive built into macroeconomics. That the elephant dead is worth more than the elephant alive is and so as the rhino and so is the – and. So, how do you have a situation where that's the nature of incentive, where you're incentivizing an activity and then trying to bind it or keep it from happening? The same would be true with over-fishing. As long as live fish are worth nothing and dead fish are worth more, and you have ... There's something fundamentally perverse about the nature of the economic incentive. The same is true that when war goes ... When we have war and there's more military manufacturing, GDP goes up. When there's more addiction and people are buying the supply of their addiction, GDP goes up. When there are more sick people paying for healthcare cost, GDP goes up.

Daniel Schmachtenberger: So, it's obviously a perverse kind of metric. So, anytime someone can fiscally advantage themselves or a corporation can, in a way that either directly causes harm or indirectly externalizes harm, we have to fundamentally solve that. If there's something like \$70 trillion a day of activity happening that is a decentralized system of incentives that is incenting people to do things that are directly or indirectly causing harm, there's really nothing we can do with some billions of dollars of nonprofit or state or whatever money that is going to solve that thing. So, we have to say, well, what changes at the level of macroeconomics need to happen where the incentive of individuals and the incentive of corporations and the incentive of nations is more well aligned with the well-being and the

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

incentive of others, and so we're less fundamentally rivalrous in the nature of our incentive?

Daniel Schmachtenberger: So, we can see that underneath heaps of the problems, structures of macroeconomic incentive are there. That's the kind of maybe the first one that most people see. We can go deeper to seeing that even as an expression, because, whether it's a economic incentive for a corporation or whether it's a power incentive, a political power incentive for a political party or for a country, they're both instantiations of rivalrous-type dynamics that end up driving arms races because if you win at a rivalrous dynamic, the other side reverse engineers your tech, figures out how to make better versions, comes back, which creates an exponentiation and warfare. And eventually, exponential warfare becomes self-terminating on a finite planet. Exponential externalities also become self-terminating. So, if we want to say, what are the underlying generator functions of catastrophic risk, first, maybe just to make clear, the catastrophic risk landscape. Is this all right if we do a brief aside on that?

Tristan Harris: Yeah. Let's do it. Then, I think we ... Let's do that, and then let's recap just what these structures are, so people are tracking each of these components because you've already mentioned a few different things. I mean, the first thing is just, many listeners might hear what you're sharing as an overwhelming set of problems. And, I think just to recap, it's important people understand that it's overwhelming if you're not using a problem solving framework that allows you to see the interconnected nature of those problems because if you solve them with the limited tools we have now, of let's just solve the social media problem by pulling one lever and changing one business model of one company or banning TikTok, but then you get 20 other TikToks that come and sit in its place with the same perverse incentive of addiction, the same rivalrous dynamic competing for human attention, we're going to end up perpetuating those problems. So, just to sort of maybe recap some of that for listeners, and I think maybe let you continue with the other generator functions. Let's just make sure that people really get those frameworks. I think it's really important.

Daniel Schmachtenberger: Yeah. I mean, in the case that you and Center for Humane Technology have brought so much attention to with regard to the attention harvesting and directing economy, it's fair to say that it probably was not Facebook or Google's goal to create the type of effects that they had. Those were unintended externalities. They were second order effects. But they were trying to solve problems, right? Let's solve the problem, if we're Google, of organizing the world's information and making better search. That seems like a pretty good thing to do. Let's solve the problem with making it freely available to everybody. Well, that seems like a pretty good thing to do. With the ad model, we can make it freely available to everyone, and let's recognize that only if we get a lot of data will our machine learning get better. So, we need to actually get everybody on this thing, so we definitely have to make it free. Then, we get this kind of recursive process.

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

Daniel Schmachtenberger: Well then, the nature of the ad model doing time on site optimization and stuff I'm not going to get into because you've addressed it so well, ends up appealing to people's existing biases rather than correcting their bias, appealing to their tribal ingroup identities rather than correcting them and appealing to limbic hijacks rather than helping people transcend them. And as a result, you end up actually breaking the social solidarity and epistemic capacity necessary for democracy. So, it's like, "Oh, let's solve the search problem. That seems like a nice thing. The side effect is we're going to destroy democracy and open societies in the process, and all those other things." Those are examples of solving a problem in a way that is externalizing harm, causing other problems that are oftentimes worse. And so ... let's just focus on the opportunity and –

Tristan Harris: Typically, this will get accounted for as, oh, this is just an unintended consequence, but there's some other generator functions, I think, we should outline. I mean, if YouTube and Google didn't personalize search results and what video to show you next, then the other guy did, and TikTok starts personalizing, they're caught in a race to the bottom of whoever personalizes more for the best limbic hijack and so just to sort of connect some of those themes together for listeners.

Daniel Schmachtenberger: So, you mentioned race to the bottom. Obviously, CHT has discussed this before, and this is a key piece of the game theoretic challenge in global coordination. The two primary ways it expresses itself is arms races and tragedy of the commons. The tragedy of the commons scenario is, if we don't over-fish that area, virgin ocean, but we can't control that someone else doesn't because how do we do enforcement if they're also a nuclear country? That's a tricky thing, right? How do you do enforcement on nuclear-equipped countries? So, us not doing it doesn't mean that the fish don't all get taken. It just means that they grow their populations and their GDP faster, which they will use rivalrously. So, we might as well do it. In fact, we might as well race to do it faster than they do. Those are the tragedy of the commons type issue. The arms race version is, if we can't ensure that they don't build AI weapons or they don't build surveillance tech and they get increased near-term power from doing so, we just have to race to get there before them. That's the arms race type thing.

Daniel Schmachtenberger: It just happens to be that while that makes sense for each agent on their own in the short term, it creates global dynamics for the whole in the long-term that's self-terminating. Because you can't run exponential externality on a finite planet. That's the tragedy of the commons, one, and you can't run exponential arms races and exponential conflict on a finite planet. So, the thing that has always made sense, which is just keep winning at the arms races has had a world where we've had lots of wars increasing in their scale and lots of environmental damage. We started desertification thousands of years ago. It's just, it has been a long, slow exponential curve that really started to pick up with the Industrial Revolution and is now really verticalizing the digital revolution and the cumulative harm of that kind of thing becomes impossible now.

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

Daniel Schmachtenberger: So, basically, with the environmental destruction, with the wars and with the class subjugation things that civilizations had in the past, pretty much anyone would say we have not been the best stewards of power and technologies increasing our power. Exponential tech means tech that makes better versions of itself. So, you get an exponent on the curve ... And we're now in a process where that's ... very, very rapid computation, gives the ability to design better systems of computation. Computation and AI applied to biological big data and protein folding gives the ability to do that on biotech and on and on, right?

Daniel Schmachtenberger: So, we could say the central question of our time is if we've been poor stewards of power for a long time and that's always caused problems, but the problems now become existential ... They become catastrophic. We can't keep doing that. How do we become adequately good stewards of exponential power in time, right? How do we develop the good decision-making processes, the wisdom necessary to be able to be stewards of that much power? I think that that's a fair way to talk about the central thing.

Daniel Schmachtenberger: Now, if it's okay, the thread we were about to get to, I think, is a good one, which was the history of catastrophic risk coming up to now, is that before World War II, catastrophic risk was actually a real part of people's experience. It was just always local. But an individual kingdom might face existential risk in a war where they would lose. So, the people faced those kinds of reality. In fact, one thing that we can see when you read books like *The Collapse of Complex Societies* by Joseph Tainter and just any study of history is that all the great civilizations don't still exist. Which means that one of the first things we can say about civilizations is that they die. They have a finite lifespan on them.

Daniel Schmachtenberger: One of the interesting things we can find is that they usually die from self-induced causes. They either over-consumed the resources and then stopped being able to meet the needs of the people through unrenovable environmental dynamics, and that's old, or they have increasing border conflicts that lead to enmity that has more arms race activity coming back at them, or they have increasing institutional decay of their internal coordination processes that leads to inability to operate quickly and those types of things.

Daniel Schmachtenberger: So, we can say that it's the ... Fundamentally, most all civilizations collapse in a way that is based on generally self-terminating dynamics. We see that even when they were overtaken by armies, oftentimes they were armies that were smaller than ones they had defended against successfully at earlier peaks in their adaptive capacity. Okay. So, catastrophic risk has been a real thing. It's just been local. And, it wasn't until World War II that we had enough technological power that catastrophic risk became a global possibility for the first time ever.

Daniel Schmachtenberger: And this is a really important thing to get because the world before World War II and the world after was different in kind so fundamentally. This is why when you study history, so much of what you're studying is history of

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

warfare, of neighboring kingdoms and neighboring empires fighting. Because the wars were fundamentally winnable, at least for some, right? They weren't winnable for all the people who died, but at least for some. With World War II and the development of the bomb became the beginning of wars that were no longer winnable. If we employed our full tech, and continued the arms race even beyond the existing tech, it's a war where win-lose becomes omni lose-lose at that particular level of power.

Daniel Schmachtenberger: So, that created the need to do something that humanity had never done, which was that the major superpowers didn't war. The whole history of the world in the history of the thing we call civilization, they always did. So, we made an entire world system, a globalized world system that was with the aim of preventing World War III. So, we could have non-kinetic wars, and we did, right? Increasingly, you can see from World War II to now, a movement to unconventional warfare, narrative and information warfare, economic, diplomatic warfare, those types of things, resource warfare. And you could, if you were going to have a physical kinetic war, it had to be a proxy war, but to have a proxy war, that also required narrative warfare to be able to create a justification for it.

Daniel Schmachtenberger: Also, to be able to prevent the war ... So, the post-World War II, Bretton Woods, mutually assured destruction, United Nations world was a solution to be able to steward that level of tech without destroying ourselves. It really was a reorganization of the world. It was a whole new advent of social technologies or social systems just like the U.S. was new social technologies or social systems coming out of the Industrial Revolution. The Industrial Revolution ended up giving rise to nation state democracies.

Daniel Schmachtenberger: The nuclear revolution in this way kind of gave rise to this IGO, inter-governmental world, and it was predicated on a few things. Mutually assured destruction was critical. Globalization and economic trade was critical, that we, if the computer that we're talking on and the phone that we talk on is made over six continents and no countries can make them on our own, we don't want to blow them up and ruin their infrastructure because we depend upon it, so let's create radical economic interdependence, so we have more economic incentive to cooperate. Makes sense. Let's grow the materials economy so fast through this globalization, that the world gets to be very positive GDP and gets to be very positive sum so that everybody can have more without having to take each other's stuff.

Daniel Schmachtenberger: That was kind of like the basis of that whole world system. We can see that we've had wars, but they've been proxy wars and Cold Wars. They haven't been major superpower wars, and they've been unconventional ones. But we haven't had a kinetic World War III. We have had increase of prosperity of certain kinds, 75 years, give or take. Now, we're at a point where that radically positive sum economy that required an exponential growth of the economy, which means of the materials economy, and it's a linear materials economy, that unrenovably takes resources from the Earth faster than they can reproduce themselves and turns them into waste

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

faster than they can process themselves, has led to the planetary boundaries issue, where it's not just climate change or over-fishing or dead zones in the ocean or microplastics or species extinction or peak phosphorus. It's a hundred things, right?

Daniel Schmachtenberger: There's all these planetary boundaries, so we can't keep doing exponential linear materials economy. That thing has come to an end because now, that drives its own set of catastrophic risks. We see that the radical interconnection of the world was good in terms of we'll not bomb each other, but it also created very high fragility because what it meant is, a failure anywhere could cascade to failures everywhere because of that much dependence. So, we can see with COVID, we had what was a local issue to an area of China but because of how interconnected the world is with travel, it became a global issue at the pandemic level. And, it also became an issue where to shut down the transmission of the virus, we shut down travel, which also means shut down supply chains, which meant so many things, right, and very fundamental things that weren't obvious to people at first like that a country's agriculture depends upon the shipment of pesticides that they don't have stored. So, we got these swarms of locusts because of not having the pesticides, which damaged the food supply and shipments of fertilizer and shipments of seed.

Daniel Schmachtenberger: So, we ended up seeing a drive of the food insecurity, of extreme poverty, at a scale of death threat that is larger than the COVID death threat was as a second order effect of our problem. We were trying to solve the problem of don't spread COVID, and the solution had these massive second, third order effects that are still playing out, right? And that was a relatively benign pandemic, a relatively benign catastrophe compared to a lot of scenarios we can model out. So, we can say, "Okay, well, we like the benefit of interconnectivity, so we're not invested in bombing each other, but we need more anti-fragility in the system."

Daniel Schmachtenberger: Then, the mutually assured destruction thing doesn't work anymore, because we don't have two countries with one catastrophe weapon that's really, really hard to make an easy to monitor because there's not that many places that have uranium. It's hard to enrich it. You can monitor it by satellites. We have lots of countries with nukes. But we also have lots of new catastrophe weapons that are not hard to make, that are not easy to monitor, that don't even take nation states to make them. So, if you have many, many actors of different kinds with many different types of catastrophe weapons, how do you do mutually assured destruction? You can't do it the same way.

Daniel Schmachtenberger: So, what we find is that the set of solutions post-World War II that kept us from blowing ourselves up with our new power lasted for a while, but those set of solutions have ended, and they have now created their own set of new problems. So, there's the catastrophic risk world before World War II, the catastrophic risk world from World War II till now, and then the new thing. So, the new thing says we have to have solutions that deal with the planetary boundary issues, that deal with global fragility issues, and

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

that deal with the exponential tech issues both in terms of the way exponential tech can be intentionally used to cause harm i.e., exponential tech-empowered warfare and unintentionally i.e., exponential tech-empowered externalities, and even just totally unanticipated types of mistakes, the Facebook, Google type problem multiplied by AGI and things like that.

Daniel Schmachtenberger: So, when we talk about what the catastrophic risk landscape is, that's the landscape. The meta crisis is how do we solve all of that. And, recognizing that our problem-solving mechanisms haven't even been able to solve the problems we've had for the last many years, let alone prevent these things. So, the central orienting question, it's like the UN has 17 Sustainable Development Goals. There's really one that must supersede them all, which is develop the capacity for global coordination that can solve global problems. If you get that one, you get all the other ones. If you don't get that one, you don't get any of the other ones. So, we can talk about how do we do that, but that becomes the central imperative for the world at this time.

Tristan Harris: So, you're saying a whole bunch of things, and one thing that comes to mind here, if I'm just reading back some of the things you've shared, the development of the, let's call it one of the first exponential technologies which is the nuclear bomb, led to a new social system, which was sort of the post-Bretton Woods world of trying to stabilize that one exponential technology in the world in a way that would not be catastrophic. And even there, we weren't able to sort of make it all work. I think people should have maybe a list of some of the other exponential technologies because I want to make sure that that phrase is defined for listeners.

Tristan Harris: There's a lot of different ways that we've now not just created more exponential technologies but more decentralized exponential technologies. I think people should see Facebook and Google as exponential attention mapping or information driving technologies that are shaping the global information flows or the wiring diagram of the sort of global societal brain at scales that are exponential. It's sort of a nuclear-scale rewiring of the human civilization. We couldn't do that with newspapers. We couldn't do that with a printing press, not at the scale speed, et cetera, that we have now. So, do you want to give maybe some more examples of exponential technologies because I think that's going to lead to, we're going to need new kinds of social systems to manage this different landscape of not just one exponential nuclear bomb but a landscape?

Daniel Schmachtenberger: Yeah. Indulge me as I tell a story first that leads into it because it'll be a relevant framework.

Daniel Schmachtenberger: Obviously, the bomb was central to World War II and the world system that came afterwards and what motivated our activity getting into it. But it was not the only tech, it was one new technology that was part of a suite of new technologies that could all be developed because of the level science had gotten to. And basically, physics and chemistry had gotten to

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

the point that we could work on a nuclear bomb. We could start to work on computation. We could get things like the V-2 rocket and rockets, and a whole host of applied chemistry.

Daniel Schmachtenberger: One way of thinking about what World War II was, it's not the only way of thinking about it, but it's a useful frame, and I think it's a fair frame, is that there were a few competing social ideologies at the time, primarily, German fascism, fascism, socialism, whatever you want to call it, Soviet communism and Western liberalism, something like that.

Daniel Schmachtenberger: This new suite of technologies, whoever developed it and was able to implement it at scale first would win. That social ideology would win because it's just so much more powerful. If you have nukes and they have guns, you're going to win, right? And Germans were actually ahead of both the US and the Soviets because of some things that they did to invest in education and tech development, but that led both the Soviets and the US to really work and to catch up as fast as they can. When the US finally figured it out, which we were actually a little bit slow to. Einstein actually wrote a letter, the Einstein-Szilard letter that went to the US government saying, "Now the science really does say that this thing could happen and the Germans could get it, and you should focus on it."

Daniel Schmachtenberger: And at first they didn't take him up on it. It wasn't until the private sector actually, non-profit supported, advanced it further, then the Manhattan project was engaged in. But then it was engaged in when they recognize the seriousness, that there was an actual eminent, existential risk to the nation and the whole Western ideology and whatever, then it was an unlimited budget, it was, "Let's find all the smartest people in the world and let's bring them here and let's organize, however we need to, to make this thing happen. And let's do it for all of the new areas of tech. We're going to get the Enigma machine and crack the Enigma code. We're going to get a V-2 rocket. We're going to figure out how to reverse engineer that and advance rocketry. We're going to do everything needed to make a nuclear bomb, and then more advanced ones."

Daniel Schmachtenberger: That was the biggest jump in technology ever in the history of the world, in recorded history, as we know it. And it wasn't actually done by the market. It was done by the state. That's a very important thing. This idea that markets innovate and states don't innovate, it's just historically not true here. This was state funds and state controlled operation, in the same way that the Apollo project coming out of it was.

Daniel Schmachtenberger: And a technological jump of that kind hasn't happened since. So it's an important thing to understand. But we can say, though this is not a totally fair thing to say, we can say that the US came out dominant in that technological race, the US and the USSR both had a lot of capacities, so that was the Cold War. And then finally the US came out. And so the post-World War II system was a US led system. The UN was in the US, the Bretton Woods system was pegged to the US dollar.

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

Daniel Schmachtenberger: What I would say is that, so it wasn't one type of tech. It was the recognition that science had got to a place where there was going to be a whole suite of new tech, and new tech meant more power. And whoever had the power would determine the next phase of the world. And if we didn't like the social ideologies that were going to be guiding it, of course, we can also think of it as just who wanted to win at the game of power, but from the philosophical argument, if we didn't like the social ideologies then we'd have another social ideology to get it.

Daniel Schmachtenberger: What I would say is that there is an emerging suite of technologies now that is much more powerful in the total level of technological jump than the World War II suite was. In fact, orders of magnitude more. And only those who are developing and employing exponential tech will have much of a say in the direction of the future, because just from a realpolitik point of view, that's where the power is. And if you don't have the power, you won't be able to oppose it.

Daniel Schmachtenberger: And so what do we mean by exponential tech? There's a couple of different ways of thinking about it. Just exponentially more powerful is a very simplistic way. And in that definition, nuclear is exponential tech. But what we typically mean with exponential tech is tech that makes it possible to make better versions of itself, so that there is a compounding interest curve. The tech makes it easier to make a better version, which makes it easier to make a better version. And so we see that starting with computation really in a fundamental way, because computation allows us to advance models of computation. How do we make better computational substrates? How do we get more transistors in a chip? How do we make better arrangements of chips so we get GPUs and those types of things?

Daniel Schmachtenberger: And so, in this new suite of technology, the center of it is computation. The very, very center of that is AI, self-learning computation on large fields of data. The other kind of software advances like various meaningful advances in cryptography and big data, and the ability get data from sensors and sensor processing, image recognition, all like that, is a part of that central suite. And the application of that to the directing of attention and the directing of behavior by directing attention, which you focused on very centrally.

Daniel Schmachtenberger: Then the next phase is the application of the tech, the application of computation to increasing computational substrate. So this is now the software advancing the hardware that can advance the total level of software, you see the recursion. So that's not just continuously better chips, it's also quantum computing, photo computing, DNA computing, those other types of things. And the other types of hardware that need to be part of that thing, i.e., sensor tech in particular, so that you can keep getting more data going into that system that can do big data machine learning on it.

Daniel Schmachtenberger: Then it's the application of that computation and AI specifically, to physical tech. So to nanotech, material sciences, biotech, and even things like

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

modeling how to do better nuclear and robotics and automation. And so when you start thinking about better computational substrates, running better software with more total data going in with better sensors in better robots, you start getting the sense of what that whole suite of things looks like.

Daniel Schmachtenberger: So that's the suite of things that I would say is what we would call exponential tech. And the reason why the term exponential is important is we don't think exponentially well, our intuitions are bad for it because we think about how much progress was made over the last five years. We imagine there'll be a similar amount over the next five years, and that's not the way exponential curves work. And so it's very hard for us. Our intuition was calibrated on the past, and it's going to be miscalibrated for forecasting the total rate of change and the magnitude of change.

Tristan Harris: So to link this for one much more narrow aspect for our listeners who are familiar with social media and the social dilemma, and you're talking about self-compounding systems that improve recursively like that. If I'm TikTok, or if I'm Facebook and I use data to figure out what's the thing to show you that's going to keep you here for longest. And it's going to bypass your prefrontal cortex and go straight to your limbic system, your lizard brain. Well, the better it gets at doing that and succeeding at that, the more data it has to make a better prediction the next time, but then a new user comes along who it's never seen before, but hey, they're clicking on exactly the same pattern of anorexia videos that we've seen these other 2 million users have that turn out to be teenage girls. And it just happens to know that this other set of videos that are more anorexia videos are also going to work really, really well.

Tristan Harris: So there's a self-compounding loop, it's learning not just from one person and getting a better version of hijacking your nervous system, but learning across individuals. And so now you get a new person coming in from some developing country who's never used TikTok before, and they're just barely walking in through the front door the very first time. It's sort of like when Coca-Cola goes to Southeast Asia for the first time and you get diabetes 10 years later, because you refined all the techniques of marketing so effectively, but now happening at scales that are automated with computation.

Tristan Harris: So what you're talking about is the impact of computation and learning on top of learning data on top of data, and then cross-referencing lookalike models and all of this kind of thing you could apply to the domain, at least that social dilemma watchers and people who are familiar with our work might be able to tie into.

Daniel Schmachtenberger: The more people you have in the system and the more data per person that you're able to harvest, the more stuff you have for the machine learning to figure out patterns on, which also means that the machine learning can provide things that the users want more, even if it's manufactured want, even if it's manufactured demand, which means that

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

then more users will come and put more data in, and it can specifically figure out how to manufacture the types of behavior that increase data collection. And so you do get this recursive process on how many people, how much data, how good are the machine learning algorithms, that kind of thing.

Daniel Schmachtenberger: This is one of the reasons that we see these natural monopoly formations within these categories of tech. And this is another reason that it's important to understand that, these types of self-reinforcing dynamics and things like network effects like Metcalfe's law, didn't exist when the Scottish enlightenment was coming up with its ideas of capitalism and market and the healthy competition in markets and why that creates checks and balances on power, they didn't exist. Adam Smith did not get to think about those things.

Daniel Schmachtenberger: And so when you have a situation where the value of the network is proportional to the square of the people coming into the network, then your incentive to keep it free upfront, maximize addition, drive behavior into the system. And then once you get to the breakaway point on the return of that thing, it becomes nearly impossible for anyone else to come in and overtake that thing. So you get a power law distribution in each vertical. You get one online market that is bigger than all the other online markets, one video player that's bigger than all the other video players. One search, one social network. That's not because of a government monopoly, that's because of this natural tech monopoly. This also means that when we created the laws around monopolies, they don't apply to this thing.

Daniel Schmachtenberger: And yet this thing still has the same spirit of power concentration and unchecked power that our ideas of monopoly had, but it's able to grow much faster than law is able to figure out how to deal with it, or faster than economic theory can change itself. And so one of the things that we see is that our social technologies like law, like governance, like economics are actually being obsoleted by the development of totally new types of behavior and mechanics that weren't part of the world they were trying to solve problems for.

Daniel Schmachtenberger: And so the Scottish Enlightenment was the development of new ideas of how to problem solve the problems of its time. The constitution was trying to figure out how to solve the problems of its time. I would say they were good thinking, they were good work. The Bretton-Woods world was. None of them are adequate to solve these problems because these problems are different in kind, and even where they're just an extension of magnitude, when you get enough change in magnitude, sometimes it becomes a difference in kind. Like, as you're getting more and more information to process, once you get past what humans can process, info singularity type issues, okay, well now it's a difference in magnitude that becomes a difference in kind, which means you need a fundamentally different approach.

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

Daniel Schmachtenberger: So I would say this is where it's important to recognize that those social technologies that we loved so much because they seem so much better than all the other options we had at the time, like markets and like democracy. These are not terminal goods in and of themselves. The terminal goods were things like human liberty and justice and checks and balances on power and opportunity and distribution of opportunity and things like that. These were the best social technologies possible at the time, the new technologies both killed those things. They don't work anymore. You can't have the social technology of the Fourth Estate that was necessary for democracy, which is why the founding fathers said things like, "If I could have perfect newspapers and a broken government or perfect government and broken newspapers, I'd take the newspapers." Because if you have an educated populace that all understands what's going on, they can make a new form of government. If you have people that have no idea what's going on, how could they possibly make good choices if their sense-making is totally broken?

Daniel Schmachtenberger: So we had this idea that the Fourth Estate was a prerequisite to a participatory governance, but that was based on a very narrow, limited capacity for print. And again, it was the technology of the Gutenberg Press that was one of the things that actually ended feudalism. And so the founding fathers were employing that new tech, both because it upended the previous tech and it made this new thing possible. Same with guns, they needed guns and the second amendment to make this new thing possible.

Daniel Schmachtenberger: But once we get to a internet world where you don't have centralized broadcast, you have decentralized, and then there's so much stuff that you can never possibly find it all in search or whoever coordinates the search, the content aggregators, which is the Facebook, YouTube, whatever, are doing it with the types of business models we have, the Fourth State is just dead forever. That old version. There's no way to recreate that version. So that either means democracy is dead forever or anything like a well-informed citizenry that could participate in its governance in any form, or you have to say, "What is a post-internet, post-social media, post-info singularity Fourth Estate that creates an adequately educated citizenry?"

Daniel Schmachtenberger: That's thinking and about the way that our social technologies, our social systems have to upgrade themselves in the presence of the tech that obsoleted the way they did work, but we can also see, and we can give examples of this, how the new tech also makes possible new things that weren't possible before. So we can do something better than industrial-era democracy or industrial-era markets. Which is why I say they aren't a terminal good, they are a way to deliver certain human values that really matter. And the new technology that obsoletes those can actually also be facilitative in designing systems that also serve those values. But it's not a given that it does, that has to become the central orienting mission.

Tristan Harris: Dan, just to make sure we're linking this back to the start of this conversation. We started this conversation by saying, the way that we're going about solving problems, let's say using the legacy systems of law

**Center for Humane Technology | Your Undivided Attention Podcast**  
**Episode 36: UNEDITED A Problem Well-Stated is Half Solved**

making in a Congress or using the legacy systems of a town hall to vote on a proposition, or trying to pass laws as fast as social media is rewiring society, the lines don't match. And so what you're saying is that, and just for listeners, because I know that you use the phrase social technology, but I think you're really talking about social systems, ways of organizing a democracy or –

Daniel Schmachtenberger: Technology in the most fundamental sense of the word, of something humans designed to facilitate certain kinds of activity or outcomes. Like language is a technology or democracy is a technology. So social systems.

Tristan Harris: Social systems. Yeah. And so if the old world approach of, some of you might be hearing this and say to themselves, "Now hold on a second. So we have all these institutions, we have all these structures, we live in a democracy and we live in a system that is working the way it does. It has its courts. It has its attorney generals. It has its litigation procedures. It has its lawmaking bodies. If you're saying that we can't use those things because they're not adequate, or they won't help us solve those problems, we need to have new social systems. Maybe you could give us some hope about why that might be feasible, instead of feeling impossible, because this is actually precedented in our history." When new technologies show up and then new social systems emerge to make room for those technologies functioning well. You briefly touched on them, but I think it's important to give listeners a few concrete examples.

Daniel Schmachtenberger: There's a number of good academics and disciplines of academics that look at the history of evolutions in physical technology, and the corresponding evolutions and thought and culture and social systems. Marvin Harris, the cultural materialism, did a major opus work here where he specifically looked at how changes in social systems and cultures followed changes in technology. There are other bodies of work that will look at the social systems as primary or the cultures as primary. And we can say they're inter-affecting. But for instance, the vast majority of human history was tribal, was however much, 200,000 years of humans in these small Dunbar number of villages, there was a social technology, social systems that mediated that, that that had to do with how the tribal circles worked and the nature of how resources were shared. It was a very different kind of economic system, a very different kind of judicial system, a different educational system.

Daniel Schmachtenberger: It had all those things. It had a way of education, meaning intergenerational knowledge transfer of the entire knowledge set that was needed for the tribe to continue operating. The development of certain technologies, particularly the plow, but baskets and a few other things obsoleted that thing, because all of a sudden, it made possible big amounts of surplus that made reason for much larger populations to emerge. Those larger populations were going to win in conflict against the smaller populations. And so you can see that then the emergence of new social technology to facilitate large groups of people, empire types, civilization technology emerged, you can see, and that there were a few other shifts in technology

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

that evolved the types of empires that were there. And then the next one that people talk about a lot is the Industrial Revolution from the printing press specifically, and then steam engine, the gunpowder revolution was part of it, that ended feudalism and began the nation state world.

Daniel Schmachtenberger: And so you can see, what is the thing that the founding fathers in the US were doing? Well, they weren't trying to keep winning at feudalism, there was a game that had been happening for a long time. And they were saying, "We're all people who are in, of the type of people who could do well at that system. And rather than do that, we recognize that there are fundamentally things wrong with this system and fundamentally new possibilities that hadn't been previously recognized. So we're going to actually try to design a fundamentally different system, a more perfect union that makes life, liberty and the pursuit of happiness better for everybody, and increases productive capacity and things like that." So that was fundamentally an advance in social technology or social systems that both utilized new, physical technology and was enabled by it.

Daniel Schmachtenberger: In the current situation, there are groups that are advancing the exponential technologies and what that means, whatever social systems that they're employing are the social systems of the future if we don't change it. And that's what I want to get to in a moment, but who is working to implement any of the new emerging tech for better social systems that are aligned with social systems we want? You've had Audrey Tang on the show. Do you want to just briefly describe an example of what she and what they have done there? If people aren't aware of it, that's a pretty prime example for this particular iteration.

Tristan Harris: Sure, and maybe just to go back briefly, because you gave this example in one of our earlier conversations that the printing press could have been used by the feudal lords for consciously reinforcing feudalism. But instead, they said, "Actually this new technology, the printing press gives way to new ways of organizing society. And we can actually have things like a Fourth Estate or newspapers or things like that." And so —

Daniel Schmachtenberger: Both happened.

Tristan Harris: Both happened. Yeah. Right. And then the new thing, theoretically, has to win out over the old thing, or at least the one that we want, that holds the values that the society wants. So, I think a lot of people can hear our conversation. We've had this riff before actually, following our last episode after my Senate testimony, speaking about a frame that you have offered and know well, which is that we can notice that digital authoritarian societies right now, like China, are consciously using exponential technologies to make stronger, more effective digital closed and authoritarian societies.

Tristan Harris: And in contrast, digital open societies, democracies like the United States are not consciously using technology to make stronger, healthier, open societies. Instead, they've surrendered what they are to private technology,

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

multinational corporations pursuing self-interest to shareholders and are profiting from the degradation and dysfunction of democracies.

Tristan Harris:

And so when we say all this, and we talk about how do we build the next social system and Audrey Tang and her work, I think people get tripped up in thinking that what we really mean is we have to make some kind of 21st century digital democracy. In fact, I've probably said those words, but what we're really talking about here is some new concept that preserves the principles of what we meant by a democracy, but instantiated with the new technologies, our version of the new printing press, which is networked information environments and all of the new capacities that we have in the 21st century with mobility where everyone's connected to everywhere and everything all at once.

Tristan Harris:

So what is that system, that new social system that leverages the current technology and makes a stronger, healthier open society? And I think Audrey Tang's work, I would probably send listeners back to listen to that episode. I think it's one of our most listened to and most popular episodes for a reason, because in Taiwan, she's essentially built an entire civic technology ecosystem in which people are really participating in the governance of their society. We need masks, we need better air quality sensors, we need to fix these potholes. There are processes by which every time you're frustrated by something, you actually get invited into a civic design process where, whether it's the potholes or the masks, you can actually participate in having a better system. You're complaining about the tax system and filing your taxes and maybe it's an inefficient form or something like that, you get brought into a design process of what would make it better.

Tristan Harris:

And so the system is participatory, but not in that 18th century way of, hey, there's a physical wooden townhouse and we're going to walk into it and we're going to hang out there for three hours and I'm going to yell and scream about issues that are more local within 10, 15 miles of where we are, because we were existing in a world before automobiles. We're now talking about how do you do an open society and social system, but in a world with all of the new technologies that are not just here today, but emerging.

Tristan Harris:

And so do you want to talk a little bit about how will we even navigate that challenge? And why is some new social system like that necessary for dealing with these problems that you've laid out at the beginning? I'm sure people would like to feel less anxiety about those things hanging around for longer.

Daniel Schmachtenberger:

I think what Taiwan has been doing and what Audrey Tang in the digital ministry position in particular has been leading is probably the best example, certainly one of the best examples in the world of this kind of process and thinking, and does it apply in the, or could it apply in the exact same way to the US? No, of course not. We know that because of the relatively small geography and high speed train transportation, you can get

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

across Taiwan in an hour and a half. And so when you're mentioning the small scale of local government at the beginning of the US where you come to the town hall, in a way they have that, it's 23 million people, but there is an older shared culture. There also happens to be an existential threat just right off their border that is big enough that they can't just chill and not focus on it. Everyone has to be civically engaged with some civic identity like that.

Daniel Schmachtenberger: They didn't start making their culture in the industrial era and then have to upgrade it. They started later, where they were able to start at a higher level of the tech stack. So there's a number of reasons why it's different. So we're not going to naively say, "What you do in a tiny country that is culturally and ethnically homogeneous and has a higher GDP and education per capita and whatever is the same thing you would do," but we can certainly take a lot of the examples and say, "How would they apply differently in different contexts?"

Daniel Schmachtenberger: So, the thing we said earlier, that this suite of exponential technologies is so much more powerful than all of the previous types of power, that only those who are developing and deploying them will be really steering the direction of the future. And that there are ways of employing them that do cause catastrophic risk. And the catastrophic risk is of two primary kinds, conflict theory mediated, and you just can't do warfare with this level of technology and this interconnected world and make it through well. Not all catastrophic risks means existential, it doesn't all mean nuclear war and nuclear winter, and we've killed all the mammals on earth. It might just mean we break global supply chains, kill lots of people and regress humanity and the quality of the biosphere pretty significantly.

Daniel Schmachtenberger: So I'm not just focused on existential risk. I'm interested in catastrophic risk at scale, in general. And we can see the exponential tech applied as in conflict theory and mistakes, as externalities and the cumulative effects —

Tristan Harris: Could you define conflict theory and mistake theory for people who are not familiar with those terms?

Daniel Schmachtenberger: Yeah. There's a very nice discussion on the Less Wrong forum, if people are interested to go deeper. And there's this question of how much of the problems in the world are the result of conflict theory versus mistake theory. Meaning, conflict theory is, we either wanted to cause that problem, that harm to whomever, as in, knowingly wanted to win at a war, or at least we knew we were going to cause that problem and didn't care because it was attached to something we wanted. Conflict theory. Or mistake theory, we didn't want to cause it and we really didn't know, and it was just unintended, unanticipatable consequence. And it's fair to say that there's both, there's plenty of both.

Daniel Schmachtenberger: One thing that is worth knowing is that, if I'm trying to do something that is actually motivated by conflict theory, it benefits me to pretend that it was mistake theory. It benefits me to pretend that I had no idea. And then

## Center for Humane Technology | Your Undivided Attention Podcast Episode 36: UNEDITED A Problem Well-Stated is Half Solved

afterwards say, "Oh, it was an unintended, unanticipatable consequence. It was too complex. People can't predict stuff like that." And so the reality of mistake theory ends up being a source of plausible deniability for conflict theory, but they're both things, and we have to overcome both. Meaning we have to have choice-making processes in our new system of coordination.

Daniel Schmachtenberger: This sounds like maybe hippy stuff until you take seriously the change of context. "Oh, we have to have problems of choice-making that consider the whole." That sounds like unrealizable hippy stuff, until you realize, but we're making choices that affect the whole, at a level that can even individually be catastrophic and is definitely catastrophic cumulatively. So if we aren't factoring it, then the human experiment self-terminates, and maybe that's the answer to the Great Filter hypothesis. And so our —

Tristan Harris: I think people don't have an intuitive grasp of what it means that each of us are walking around with the power of gods to influence huge, enormous consequences. I could give a few examples, every time you enact with a global supply chain and hit buy on Amazon, you invisibly enacted shipping and planes and petroleum and wars in the middle east. There's a whole bunch of things that we're tied into. When you are posting something on social media and have more than a million followers, you're influencing a global information ecology. And if you're angry and biased about one side or the other of the pandemic is real, or it's not real, or something like that, you're externalizing more bias into the comments of how the rest of the world understands things.

Tristan Harris: So we're walking around with increasing power, but I don't think the increasing power that we've granted is as intuitive for some folks. Could you explain some more examples of that?

Daniel Schmachtenberger: There's both cumulative effect and, cumulative long-term and fairly singular short term. And, cumulative long-term, you go back to early US settlers coming into the US, moving west and there being buffalo everywhere. And there had been buffalo everywhere for a very long time. And then there's no buffalo. And whole areas that were forested with old growth forests became deforested. And it was like, "No, it's impossible. We could never get rid of all the buffalo, we could never cut down all the trees," but the cumulative effect of lots of people thinking that way were individually, "I have no incentive to leave the buffalo alive. And I do have an incentive for my family individually to kill it." But everybody thinking that way and increasing our desire for how much we consume per capita, our technology that allows us to

Daniel Schmachtenberger: consume more per capita and developing more capita, more total people. Well, now then you start getting environmental destruction and species extinction at scale. And that's a long time ago, right, that's much lower tech and much less people. And it's distributed action, it's a cumulative effect issue and obviously nobody's intending to fill the ocean with microplastics. But everybody's buying shit that is filling the oceans with microplastics. And

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

so everyone is participating with the system where the system as a whole is sociopathic. The system is self-terminating. The system doesn't exist without all the agents interacting with it. All the agents feel like their behavior is so small that that justifies everybody doing that thing, right? So that's what we mean by cumulative, catastrophic risk.

Daniel Schmachtenberger:

But it's also true that whoever made that thermite bomb and hooked it to a drone and hit the Ukrainian munitions factory a couple of years ago that caused a billion dollars in damage, exploded the ammunition's factory. The effect of a bomb as big as the largest non-nuclear bomb the US arsenal has, an incendiary bomb. That was a homemade little bomb in a drone, right? And CRISPR gene drives are cheap and easy, and it doesn't take that much advanced knowledge to start working with them. And so, that starts to look like individuals and small groups with real catastrophic ability, not long-term and cumulatively. The increase in our tech gives us both issues, via globalization and the overall system, you get these cumulative long-term effects. And with the exponential tech creating decentralized catastrophic capabilities, one of the core questions we have to answer is how do we make a world that is anti-fragile in the presence of those kinds of catastrophic capabilities that are easy to produce and thus decentralizable.

Tristan Harris:

And so how do we do that? What are the social systems that we need to employ to bind some of these bad effects in ways that the natural inclinations of self-interested actors will drive things in that direction?

Tristan Harris:

Just to link this to the social media space for people, if I know that I can get a little bit more attention and a little bit more likes, and clicks, and follows, and shares and so on. If I exaggerate the truth by 5%, just to use a little bit more of an extreme adjective, I know that in the long run would be bad if everybody did that, but for me right now, I can win a few hits and I can get more influence. And I'm an Instagram influencer and I'm making \$10,000 a month, and if I don't do it, I'm noticing everyone else is doing it. And if I don't use the filter, everyone else is using the filter. And so, everyone ends up in this race to the bottom situation that has that cumulative degradation or cumulative derangement, where there's increasing distance between what is true and what people believe, because we've all been subtly exaggerating it to make our point and gain influence and so on. And so, just to give another example, maybe for listeners in the space that they're more familiar with.

Tristan Harris:

But going back, I mean, the whole premise of this is as we gain more exponential technologies that have more capacity and more hands — so instead of having just the US and Russia having this, you have as you mentioned, CRISPR gene drives or some of the drone things that are out there, more and more people have access to these things. How can we

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

bind those kinds of forces, and what are the social systems that we need to make that happen?

Daniel Schmachtenberger: Yeah. I want to go back as you were describing this, I was thinking about how many people who listen to your show who maybe work in technology, who might have, they work in technology because they see the positive things technology can do and have more of a techno-optimist point of view, and this overall conversation might sound very techno pessimist. And like, did we not read Pinker and watch Hans Rosling and those types of things? And so, I want to speak to that briefly. First, this is a meta-point but it's worth saying right now, particularly on this podcast and in the post-truth or fake fact world, where then so much of the emphasis has gone into, "We need fact checkers and we need real facts."

Daniel Schmachtenberger: Obviously, it's possible to have an epistemic error or even intentional error in the process of generating a fact. Is there corruption in the institutions? And that kind of thing. But let's even say that wasn't an issue and the things that go through the right epistemic process as facts are facts. Can you lie with facts? Totally. Can you mislead with facts? Yeah. Because nobody's going to make their choice on one fact, they make their choice based on a situational assessment, based on a narrative, based on a gestalt of a whole thing that's lots of different facts. Well, which facts do I include and which facts do I not include? And am I decontextualizing the fact?

Daniel Schmachtenberger: So, the quality of life has gone up so much because the average person lived on less than a dollar a day in the US in 1815, and now they live on this many dollars a day which inflation adjusted, it means a higher quality of life. Yeah, but in 1815, most of their needs didn't come through dollars. They grew their own vegetables, they hunted. So I'm decontextualizing the facts to compare something that's really apples and oranges. So even if the fact is "true," the decontextualization and recontextualization makes it seem like it means something different than it means.

Daniel Schmachtenberger: And the same with the cherry picking of facts. And I can very easily say, "Oh, there's a lower percentage of people in extreme poverty," but I might also be changing the definitions of extreme poverty. I can also, rather than focus on percentage say, "Well, there's more total people in poverty than there were total people in the world before the industrial revolution." So there's the ability to decontextualize and recontextualize facts, there's the ability to cherry pick facts. And there's the ability to Lakoff-frame facts and put particular kinds of sentiment and moral valence on it. And so, am I talking about them as illegal aliens or undocumented workers? And I get a very different sentiment. So-

Tristan Harris: Talking about it as a pre-owned car or a used car. Everyone loves a pre-owned car, no one wants a used car.

Daniel Schmachtenberger: And so, these very simple semantic frames, contextual frames, cherry picking of the things means that I can make a narrative where all the facts went through the most rigorous fact checker and yet the narrative as a

**Center for Humane Technology | Your Undivided Attention Podcast**  
**Episode 36: UNEDITED A Problem Well-Stated is Half Solved**

whole is misleading. And so, fact checking is necessary, but it is not sufficient for a good epistemics and good sense-making, and not only is it not sufficient, it's even weaponizable. This is a very important thing to understand because if you're not recognizing that, you might be believing nonsense thinking that you're using epistemic rigor.

Daniel Schmachtenberger: Okay. So the techno pessimist and the techno optimist both cherry pick, and they both Lakoff frame. And this is true with the difference in almost every political ideology, the woke and the anti-woke, the pro-socialist, pro-capitalist. You'll notice that the way they do their arguments, "The systemic racism is really, really terrible." "No, it's not that bad, the systemic racism." They both have stats, but you can almost think of it as statistical warfare, as a tool of narrative warfare. And so, this is where a higher level of earnestness, rather than a particular vested interest or bias, a higher willingness to look at biases, a higher level of rigor ends up being critical to actually overcoming these things. So can I cherry pick stats that make it look like everything is getting better? Totally. And those things are true.

Daniel Schmachtenberger: And nobody wants to go back to a world before novocaine when you have to do dentistry. And nobody wants to go back to a world before penicillin, when basic bacterial infections were around and there's totally good stuff that has emerged. And are there all kinds of ubiquitous mental illnesses and chronic complex disease that didn't exist before and increase in the total number of addictive type behaviors within populations, and a radical increase in the catastrophic risk landscape and negative effect to environmental metrics.

Daniel Schmachtenberger: So things are getting better and things are getting worse at the same time. It's important to understand that depending upon what you pick. It's just that the things that are getting worse are heading towards tipping points that make the whole thing no longer viable. And so, we're not denying that there are things that are getting better. We're saying that for the game to continue at all right, to have it be an infinite game that gets to keep continuing, there are certain things that have to not happen, and you can't have the things that are getting worse keep getting worse at the curve that they are, and have the things that are getting better be able to continue at all. So I just want to say that.

Daniel Schmachtenberger: So naive techno optimism can actually make you a part of the problem, because then you do things like develop a solution to a narrowly defined problem and externalize harm to other areas because you weren't taking seriously enough not doing that. But techno pessimism also makes you a part of the problem, or at least not a part of the solution because the future is not going to be determined by Luddites, it's not going to be determined by people who aren't developing the tools of power. So if you aren't actually looking at, how do we develop a high tech world that is also a fundamentally desirable – in terms of a high nature and high touch – world, then you really aren't thinking about it in a way that ends up mattering.

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

Daniel Schmachtenberger: And so we are techno optimist, but not naive techno optimist. We go through the totally cynical phase of, "Man, tech is a serious issue." And then you go to a post-cynical phase of, "If I want to be techno optimist and not be silly, what does it take to imagine a world where humans have that much power and we are good stewards of it?" Meaning that we actually tend to each other well and we don't create a dystopic world that has exponential wealth inequality and an underclass that nobody in the upper-class would want to trade places with, and that doesn't cause catastrophic risk. Right now, the amount of power of exponential tech makes two attractors most likely: catastrophic risk of some kind. Or social systems that do not preserve the values that we care most about that are the ones that are currently most working to develop and deploy that technology.

Daniel Schmachtenberger: And to just give a very brief recap of the frame that, Tristan, you gave earlier. As you mentioned, China is not leaving a hundred percent of its technology development to the market to develop however it wants even if it harms the nation state, they are happy to buy into technology companies that are getting too large and in ways that would damage the nation state as we saw with Ant corporation. And they are doing a lot of very centralized innovation as well associated with long-term planning. Long-term planning is a key thing. In the US, term limits make long-term planning very hard, as does a highly rivalrous two party system that is willing to damage the nation as a whole to drive party wins.

Daniel Schmachtenberger: So in that system, almost all the energy just goes into trying to win, right? You spend at least a couple years, but even the years before that of fundraising, creating alliances to just try to win. Then you're not going to invest in anything heavily that has return times longer than four years because it won't get you reelected, so no real long-term planning. And then whatever you do in those four years will get undone systematically in the next four years for the most part. All right, that system of governance will just fail comprehensively in relationship to a system that doesn't have that much internal infighting and that has the capacity to do long-term planning.

Daniel Schmachtenberger: And there's a million examples we can look at, but just when did high-speed trains start? We saw them emerge in Europe, we saw them emerge in Japan and in China. We've seen China now start to export them all around the world and the US still doesn't have any high-speed trains. And it's like, what happened? Why? And we can see that the US innovated in fundamental tech in the Manhattan project through to the Apollo project, but then it started to privatize almost everything to the market. The market started to develop in ways that really were not advancing the technology in a way that increased the coherence of the nation and the fundamental of civic values and ideas of the nation. Even the World War II thing, we can see we increased our military capacities radically, but that didn't mean we actually really advanced the ideas of democracy or those values of, do we make a better system to educate the people, and inform them and help them participate in their governance? Do we make better governance.

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

Daniel Schmachtenberger: This is why the US military is so powerful but the US government is so inept, and which is why nobody wants to fight a war with the US, a kinetic war but it's very easy right now to engage in supporting narrative warfare, where you turn the left and the right against each other increasingly, and where you do long-term planning where the US can't do long-term planning of those kinds. And so, we can see that the government of the US, and not just the US, but we can see that open societies are not innovating in how to be better open societies for the most part, more effective ones where they're using the new tech to make better open societies, that's happening in the market sector. The market is making exponentially more powerful companies. A company is not a democracy, it's not a participatory governance structure in general, it's a kind of very top-down autocratic type system.

Daniel Schmachtenberger: And so, we see that there's more authoritarian nation states that are intentionally doing long-term planning of the development in deployment of exponential tech to make better nation states of that kind. And we can't even blame them when they look at, I mean, China have the benefit of getting to see both where the USA failed, and where the USSR failed, and tried to make something that didn't fail in either of those ways and there's some things that are very smart about those approaches. So we see the exponentially empowered, more autocratic type structures, and the emergence of one natural monopoly per tech sector and then the interaction of those that kind of becomes like oligarchic feudalism, tech feudalism. Neither of those have the types of jurisprudence, or public accountability or whatever that we're really interested in.

Daniel Schmachtenberger: So the two attractors right now is, the emergence of social systems that are deploying the exponential tech, that will probably not preserve the social values that we're interested in and not be maximally desirable civilizations, probably pretty dystopic ones. Or not even guiding it well enough to prevent catastrophic risk, and catastrophic risk. Those are the two major types of attractors. We want a new attractor, which is how do we utilize the new exponential technologies, the whole suite of them, to build new systems of collective intelligence, new, better systems of social technology? How do you make a fourth estate that can really adequately educate everyone in a post-Facebook world? Well, the same way that we're trying to optimize control patterns of human behavior for market purposes, to get them to buy certain things and to direct their attention, could that be used educationally? Of course it could, if it was being developed for that purpose.

Daniel Schmachtenberger: And the AI tech that can take a bunch of faces and make a new face that is merged out of those, could it take semantic fields of people's propositions, and values and create a proposition that is the semantic center of the space? And could we use, we can't all fit into a town hall, but can we engage in digital spaces where we can have better processes of proposing refinements to the propositions? Of course we can. Could we use blockchain and other types of incorruptible ledgers to solve corruption, which is something that universally, everybody thinks is a good idea? Should

**Center for Humane Technology | Your Undivided Attention Podcast**  
**Episode 36: UNEDITED A Problem Well-Stated is Half Solved**

all government money be on a blockchain, the movement of it? So you have provenance. So you can see where the money is actually going. And if someone wants to be a private contractor, they have to agree that the accounting system, if they want government money, goes on the blockchain. So we can see the entire provenance of the taxpayer money, so that you can't have representation if there isn't transparency of how it happens.

Daniel Schmachtenberger: So when you start to think about attention-directing technology and what its pedagogical applications could be, when you start to think about AI and how it could actually help proposition development and parsing huge amounts of information to make a better epistemic commons. When you start to think about blockchain, and could we actually resolve corruption using incorruptible ledgers and making the provenance of physical supply chains, and information and money all flow across those, totally new possibilities start to emerge that never emerged before, that were never possible before. But if it doesn't become our central design imperative to develop those, those are not the highest-marketed opportunities for those right now, the highest market opportunity for blockchain is speculative tokens that have no real utility. And for AI, it is things that actually drive ads and purchasing and on and on, and for attention tech, it is the same thing.

Tristan Harris: So you've sold me on the idea that we have two dystopian attractors that we don't want. And the third attractor that we're trying to develop here is some open society that is consciously using all the modern technologies towards the values that we care about. Can you give some concrete examples of what it would look like to use AI, and attention driving tech, and click driving tech, and blockchains and all these things, but in a way that would make a stronger, healthier open society?

Daniel Schmachtenberger: Yeah, totally. So let's say we take the attention tech you've looked at so much, that when it is applied for a commercial application is seeking to gather data to both maximize time onsite and maximize engagement with certain kinds of ads and whatever. That's obviously the ability to direct human behavior, and direct human feeling and thought, in a way that is both emerged out of capitalism and has become almost a new macro-economic structure more powerful than capitalism, because even more powerful than being able to incent people's behavior with money is being able to direct what they think and feel to where the thing that they think of as their own intrinsic motive has actually been influenced or captured.

Daniel Schmachtenberger: So if we wanted to apply that type of technology, and we figured out how to make the kind of transparency that made institutions that were trustworthy enough that we could trust them with this, and already we have institutions that have it that we have no basis to trust with it, could that same tech be used educationally to be able to personalize education to the learning style of a kid or to an adult, to their particular areas of interest and to be able to, not use the ability to control them for game theoretic purposes, but use the ability to influence them to even help them learn

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

what makes their own locus of action more internalized, right? We could teach people with that kind of tech how to notice their own bias, how to notice their own emotional behaviors, how to notice group think type dynamics, how to understand propaganda and media literacy.

Daniel Schmachtenberger: So could we actually use those tools to increase people's immune system against bad actors' use of those tools? Totally. Could we use them pedagogically in general to be able to identify, rather than manufacturing desires in people or appealing to the lowest angels of their nature because addiction is profitable, can you appeal to the highest angels in people's nature, but that are aligned with intrinsic incentives and be able to create customized educational programs that are based on what each person is actually innately, intrinsically motivated by, but that are their higher innate motivators?

Daniel Schmachtenberger: Everybody can have a reward circuit that is based on chocolate cake and sloth. But the immediate spike that comes from the chocolate cake ends up then having a crash, and increased weight, and inflammation and whatever where the baseline of their happiness goes down over time. Even though every time they eat the chocolate cake, they get a spike. The exercise reward circuit is maybe not that fun, maybe even painful and dreadful in the moment, but then creates a higher baseline of energy, and capacity, and endurance, and self-esteem and you start to actually have the process become more fun. You get a new reward circuit and the baseline goes up.

Daniel Schmachtenberger: So of course I can appeal to the lower reward circuit and say, "Hey, I'm just giving people what they want." Yeah, but if you have a billion dollars or a trillion dollar organization that is preying upon the, and you discuss this very well all the time, the vulnerabilities that make people's life worse to then have the plausible deniability to say, "Yeah, but they wanted it." "Yeah, but it was a manufactured demand and a vulnerability." Where's the noblesse oblige? Where's the obligation of having that much power to actually be a good steward of power, a steward of that for other people where if there are rewards circuits that decrease the quality of their life reward circuits it increase, that we're trying to appeal to one rather than the other. Could we do that? Yeah, totally, we could.

Daniel Schmachtenberger: Could we have an education system as a result that was identifying innate aptitudes, innate interests of everyone and facilitating their development? So not only did they become good at something, but they became increasingly more intrinsically motivated, fascinated and passionate by life, which also meant continuously better at the thing. Well, in a world of increasing technological automation coming up, both robotic and AI automation, where so many of the jobs were about to be obsoleted, our economy and our education system have to radically change to deal with that. Because one of the core things an economy has been trying to do forever was deal with the need that a society had for a labor force. And there were these jobs that society needed to get done that nobody would really want to do. So either the State has to force them to do it, or you

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

have to make it to where the people who also need the jobs, so there's asymmetry, and so, the market forces them to do it.

Daniel Schmachtenberger: Well, when you technologically automate those jobs, and it happens to be that the things that are the most rote are the least fun for people and the easiest to program machines to do. And so, if you keep the same economy where if people don't produce, they don't have any basic needs met, then people want those crappy the jobs, right? But if you make it to where they have other opportunities, then of course having those jobs be automated is fine. But what does it mean to really be able to have other better opportunities? So, if one of the fundamental axioms of all of our economic theories is that we need to figure out how to incent a labor force to do things that nobody wants to do, an emerging technological automation starts to debase that. That means we have to rethink economics from scratch, because we don't have to do that thing anymore.

Daniel Schmachtenberger: So maybe if now the jobs don't need the people, can we remake a new economic system where the people don't need the jobs? Can we start to create commonwealth resources that everyone has access to? Where people's access isn't based on possession that automatically limits everyone else's access? If you get around transportation-wise with a car, based on owning that car where the vast majority of the life of the car is just sitting and not being used, for you to have access to the car, you have to have possession of it, which means that it's a mostly underutilized asset. I don't have access to the thing that you possess. Now, what we see with Uber, of course, is a situation where your access is not mediated by your possession.

Daniel Schmachtenberger: So now turn that into electric self-driving cars and now make the entire thing on a blockchain so you disintermediate even the central business, make it a commonwealth resource, and everyone has access to transportation as a commonwealth resource, it'll take a 20th of the number of cars to meet the same level of convenience during peak demand times, so much less environmental harm. It'll actually be more convenient, because I don't have to be engaged in driving the thing and there's less traffic because of the coordination and better maintenance. And there isn't an incentive for designed obsolescence in that system. You can see a situation where, "Okay, can we make it to where the wealth-augmenting capacity of that technologic automation goes back into a commonwealth because we don't have to have the same axioms of needing to incent the people?" "Oh yeah, but if you don't incent the people, they'll all be lazy welfare people." Nonsense.

Daniel Schmachtenberger: Einstein didn't do what he did based on economic incentive. And neither did Mozart and neither did Gandhi. None of the people that we are most inspired by through history were doing that. And what kids will spend so much time doing where they ask questions about why this, why this, why this, and building forts and whatever, is intrinsic motive. It's just we don't facilitate the things that they're interested in, we try to force them to be interested in things they aren't interested in. That's what ends up breaking

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

their interest in life and then they just want to hyper normal stimuli and play video games or whatever.

Daniel Schmachtenberger: What if you had a system that was facilitating their interests the entire time? Now you can have a situation where you can start to decrease the total amount of extrinsic incentive in the system as a whole. Use the automation to decrease the need for extrinsic incentive, and make an educational system and culture that's about optimizing intrinsic incentive, because if my needs are already met getting stuff and everybody's needs are met through access to commonwealth resources, there's no real status conferred. There's only status conferred by what I create. So now there is, any status is bound to a creative imperative. That's an example.

Daniel Schmachtenberger: We can look at blockchain tech even more near term and say, just to come back to this technological automation thing. So obviously it makes possible changing economics and changing education, but also, what is the role of humans in a post-AI, robotic automation world, because that is coming very, very soon. And what is the future of education where you don't have to prepare people to be things that you can just program computers to be? Well, the role of education has to be based on what is the role of people in that world. That is such a deep redesign of civilization, because the tech is changing the possibility set that deeply. So at the heart of this are deep existential questions of what is a meaningful human life? And then what is a good civilization that increases the possibility space of that for everybody? And how do we design that thing?

Daniel Schmachtenberger: We come back to blockchain and we say, "Well, blockchain is an incorruptible ledger." Well, one thing that the left and right, and everybody agrees on is that corruption happens and it's bad for the society as a whole and we don't like it. We just disagree on who does it. Is it possible that that tech could make possible decreasing corruption as a whole, it actually decreases the possibility set for corruption? Yeah. In order to do corruption, I have to be able to hide that I did it, right, I either have to break enforcement or break accounting, and mostly it's break accounting. And so, what if all government spending was on a blockchain? And it doesn't have to be a blockchain, it has to be an incorruptible ledger of some kind, Holochain is a good example that is pioneering another way of doing it, but incorruptible ledger of some kind where you actually see where all taxpayer money goes and you see how it was utilized. The entire thing can have independent auditing agencies and the public can transparently be engaged in the auditing of it.

Daniel Schmachtenberger: And, if the government is going to privately contract a corporation, the corporation agrees that if they want that government money, the blockchain accounting has to extend into the corporation. So there can't be very, very bloated corruption. Everybody got to see that when Elon made SpaceX, all of a sudden he was making rockets for a hundredth to a thousandth of the price that Lockheed or Boeing were, who had just had these almost monopolistic government contracts for a long time.

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

Daniel Schmachtenberger: Well, if the taxpayer money is going to the government, is going to an external private contractor who's making the things for a hundred to a thousand times more than it costs, we get this false dichotomy sold to us that, either we have to pay more taxes to have better national security, or if we want to cut taxes, we're going to have less national security. What about just having less gruesome bloat because you have better accounting and we make the rockets for a hundredth of the price, and we have better national security, and better social services and less taxes? Well everyone would vote for that, right, who wouldn't vote for that thing? Well, that wasn't possible before incorruptible ledgers.

Daniel Schmachtenberger: Now that incorruptible ledger also means you can have provenance on supply chains to make the supply chains closed loop so that you can see that all the new stuff is being made from old stuff, and you can see where all the pollution is going and you can see who did it, which means you can now internalize the externalities rigorously. And nobody can destroy those emails or burn those files, right? What if the changes in law and the decision-making processes also followed a blockchain process where there was a provenance on the input of information? Well, that would also be a very meaningful thing to be able to follow. So this is an example of, can we actually structurally remove the capacity for corruption by technology that makes corruption much, much, much harder, that forces types of transparency on auditability?

Daniel Schmachtenberger: What if also you're able to record history, you're able to record the events that are occurring in a blockchain that's incorruptible, where you can't change history later. So you actually get the possibility of real justice in real history and multiple different simultaneous timelines that are happening. That's humongous in terms of what it does. What if you can have an open data platform and an open science platform where someone doesn't get to cherry pick which data they include in their peer reviewed paper later, we get to see all of the data that was happening. We solve the Oracle issues that are associated, and then if we find out that a particular piece of

Daniel Schmachtenberger: science was wrong later, we can see downstream everything that used that output as an input and automatically flag what things need to change. That's so powerful. The least interesting example of blockchain is currency creation ... The capacity for the right types of accounting means the right type of choice making. Let's take AI. With AI, we can make super terrible deep fakes and destroy the epistemic commons using them, using other things like that. But we can see the way that the AI makes the deep fake by being able to take enough different images of the person's face and movements that it can generate new ones. We can see where it can generate totally new faces, averaging faces together.

Daniel Schmachtenberger: Somebody sent me some new work that they were just doing on this the other day, I found very interesting. They said, we're going to take a very similar type of tech and apply it to semantic fields where we can take everybody's sentiment on a topic and actually generate a proposition that is at the semantic center, or take everybody's sentiment and abstract from it

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

the values that they care about and create values taxonomies and say, 'We should come up with a proposition that meets all these values.' Then can you have digital processes where you can't fit everybody into a town hall, but everybody who wants to can participate in a digital space, that rather than vote yes or no on a proposition that was made by a special interest group where we didn't have a say in the proposition or even the values it was seeking to serve, so it was made in a very narrow way that like we've mentioned earlier, benefits one thing and harms something else, which is why almost every proposition gets about half of the vote and inherently polarizes the population.

Daniel Schmachtenberger: Well, people are so dumb and so rivalrous. The process of voting with bad propositions and bad representation process is inherently polarizing and downgrading to people. So what if there's a process by which there's a decision that wants to be made. You start by identifying what are the values everybody cares about. And then we say the first proposition that meets all these values well becomes the thing that we vote on. And then instead of just a direct vote, do we engage types of qualified and liquid democracy together, where you have to show that you understand the basics of that topic to be able to vote on it, but the education is free and you can keep retesting, and the basics don't show leaning one way or the other, just shows you understand the stated pros and cons so that massive populism doesn't happen. But if you don't want to come to understand it, you can cede your vote to someone else who has passed that thing.

Daniel Schmachtenberger: That type of liquid democracy, that type of qualified, educated democracy, where it doesn't have to be educated across everything, it can be per issue, and where you're not just voting on the thing, you're helping craft the propositions, these completely change the possibility space of social technology. And we could go on and on in terms of examples, but these are ways that the same type of new emergent physical tech that can destroy the epistemic commons and create autocracies and create catastrophic risks could also be used to realize a much more pro-topic world.

Tristan Harris: I love so many of those examples and especially on the blockchain and corruption one, because I think, as you said, something that the left and the right can both agree on is that our systems are not really functional and there's definitely corruption and defection going on, and just to add your example, imagine if citizens could even earn money by spotting inefficiencies or corruption in that transparent ledger so that we actually have a system that is actually profiting by getting more and more efficient over time and actually better serving the needs of the people and having less and less corruption, and so there's actually more trust and faith, and that's actually a kind of digital society that when you look at, let's say, the closed, China's digital authoritarian society, and you look at this open one, that's actually operating more for the people with more transparency, with more efficiencies, you get more SpaceX, Elon Musk type cheap ways of sending rockets to the moon and becoming a multi-planetary civilization, as

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

opposed to more bloat and more mega-monopolies, defense contractors that are not taking us to where we need to go.

Tristan Harris: That's just an inspiring vision. And I hope people listen to what you shared and go back, because there's a lot of different aspects there. I think the question on many people's minds right now is going to be, how do we get from where we are to the world that you're talking about? What are the steps that are in between?

Daniel Schmachtenberger: Obviously, I don't know. Nobody knows. Which projects emerge first and start really making success, there's a lot of different possible paths. I can say some of the things that could happen and some of the things that I think need to happen. So we take all the catastrophic risks that exponential tech makes possible and the dystopic attractors, and we say, "Okay, so we need to solve all those problems." But we're not doing really good at solving those problems right now, so our problem solving processes need upgraded, and that means new institutions. And when we say institution, we usually think of a pretty centralized thing, and with things like decentralized governance emerging, the institution might be a decentralized one, but it's. Individual people aren't going to solve all of that. So it's new institutions, centralized and decentralized that have the right capacities to solve these types of problems need to come about.

Daniel Schmachtenberger: All right, well who develops those institutions and who empowers them? And this is where the democratic idea of the power of government coming from the consent of the governed is one of the key ideas to what we would think of as the values of an open society. Let's say that there's a small number of people who think we understand these problems, we understand the solutions that must happen, everybody else doesn't get it. So we're going to make this thing happen, and because we have the power, we can just kind of implement it by force. So that becomes its own dystopia. And implemented by force might be, "Well, the people think they need to be free, so we'll implement it by attention hijacking them so that they participate with it, or don't even realize that it's happening and they just keep doing whatever's next."

Daniel Schmachtenberger: The cultural element, why we talk about the need for a new cultural enlightenment is, of course, when we look at like the founding of the US we can see all that was super wrong with it. I mean just to mention how, when Churchill said democracy is the worst form of government ever, save for all the other forms, when Socrates talked about in The Republic, when Plato was discussing it, why democracy was a dreadful idea, the arguments are good arguments. Do you want people who understand seafaring to man the boat or just the general population who knows nothing about it to man the boat? Well that's not a very good idea. Do you want the general population that knows nothing about it to build the NASA rocket? Or do you want people that know what they're doing? Well, why would we think people who have no idea what they're doing are going to be good at figuring out how a civilization should be run? What should our nuclear first

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

strike policy should be? How should we deal with the stability of the energy grid against Carrington events?

- Daniel Schmachtenberger: What does it take to have a population educated enough? And then if we say, okay, but then the other problem is if we say the people are too uneducated and maybe too irrational and rivalrous to be able to hold that power so it needs to be held by some, how do we ensure non-corruption, and who is a trustworthy authority to be able to hold that power and not have vested interest mess it up? So this is why I think it was a Jefferson quote of, the ultimate depository of the power must be the people, and if we think the people too uneducated, not enlightened to be able to hold that power, we must do everything we can to seek to educate and enlighten them, not think that there is any other safe depository.
- Daniel Schmachtenberger: So even with that, we take the US formation and you've got some founders who had read most of the books of the time, read most of the books of philosophy, knew the history of the Magna Carta and the Treaty of the Forest and all these kinds of things. Thought and talked deeply. Spent many years. Were willing to die fighting revolutionary war. Were not going along with winning at the current system, but really trying to do a fundamentally different thing to develop a new system. Not everybody who was participating in the US was doing that thing. They weren't all doing systems architecture. But they were all basically saying, we agree to that this kind of systems architecture. And we want to learn how to participate with it adequately. We'll read a newspaper. We will do a jury duty. We'll come to the town hall. That kind of thing.
- Daniel Schmachtenberger: So in Taiwan's example, I think their population is 23 million people and their online citizen engagement platform has something like 5 million people engaging. That's pretty awesome. That's not everybody, and no one should be forced to be engaging, and one of the critical things when we think deeper about is it a democracy, is that a republic, is it an epistocracy, is it... We want to think about the values, not the previous frames for them, and the values exist in dialectics, and we need to be able to hold those together. Of course you want individual liberty, but we don't want individual Liberty that gets to harm other people and other things, so we want also law, justice, collective integrity. How do you relate those things? One of the core things is the relationship between rights and responsibilities. So if I have rights and I don't have responsibilities, there ends up being tyranny and entitlement. And we can see that that's kind of rampant, the entitlement thing. If I have responsibilities and I don't have any attendant rights, it's servitude. Neither of those involve a healthy, just society.
- Daniel Schmachtenberger: So if I want the right to drive a car, the responsibility to do the driver's education and actually learn how to drive a car safely is important. And we can see that some countries have less car accidents than others associated with better drivers education. So increasing the responsibility is a good thing. We can see that some countries have way less gun violence than others, even factoring a similar per capita amount of guns, based on more

**Center for Humane Technology | Your Undivided Attention Podcast**  
**Episode 36: UNEDITED A Problem Well-States is Half Solved**

training associated with guns and mental health and things like that. So if I have a right to bear arms, do I also have a responsibility to be part of a well-organized militia, train with them, and be willing to actually sacrifice myself to protect the whole or sign up for a thing to do that? Do I have to be a reservist of some kind? Those are the right responsibility pairing.

Daniel Schmachtenberger: If I want the right to vote, is there a responsibility to be educated about the issue? Yes. Now, does that make it very unequal? No, because the capacity to get educated has to be something that the society invests in making possible for everyone. And of course we would all be silly to not be dubious factoring the previous history of these things, but this is what we then have to insist upon because do we want people who really don't understand the issues but think they do, voting? No, that's a dreadful system. But do we want people who know that something to have no avenue, or who care, do we want people who know something to have no avenue to input that into the system or people who care to have no opportunity to learn? That's also dreadful. So how do we make the on-ramps to learning available for everyone? Not enforced, but we're actually incentivizing. Can we use those same kind of social media behavior-incenting technologies to increase everyone's desire for more rights and attendant responsibilities, so that there's actually a gradient of civic virtue and civic engagement? Yeah, we can totally do that.

Daniel Schmachtenberger: So this is where the cultural enlightenment layer is. Of course not everyone is going to be working on how do we develop AI and blockchain for these purposes, but they can certainly be saying, "I am going to make sure that my representatives are talking about these issues. I want all the presidential candidates to be talking about these issues. I'm going to pay attention to and support candidates who really do in earnest ways. I'm going to invest in companies that are doing those things. I'm going to divest from companies that are doing the other things." There is a cultural enlightenment that is needed to be able to create the demand and the support for, where those projects that are earnestly working on and have the capability start to emerge.

Tristan Harris: So you've painted a compelling vision of some of the ways that a open society could consciously employ some of these technologies to revisit and re-fulfill some of the original values for which they were intended. How does this work with the existing institutions that we have? How much of this is going to rely on transforming the existing digital leviathans into something new? How much is going to depend on blockchain projects? How much of this is going to depend on existing institutions, be they the Brookings Institution or the New York Times? Can you speak to the role of new and future institutions in making this transition possible?

Daniel Schmachtenberger: Yeah. It's interesting. When we look at institutions that emerged to try to solve some social or environmental problems, or nonprofits in particular and some government branches that are associated with that, there's this kind of structural perverse incentive that, if I am an organization, which means I'm people in an organization, that have job security and some actual

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

power and access and whatever, because of this position, and my job is to solve a problem, if I fully solved the problem, I would obsolete my job and obsolete myself, so then there's this kind of perverse incentive to continue managing the problem, continue manufacturing the narrative that we're needed to manage the problem, continue manufacturing the narrative that the problem is really hard and is hard to solve, so we've got to keep doing this thing. So one of the fundamental dispositions of systems is that they want to keep existing. And yet they might no longer be fit for a purpose. They might even be antithetical to the purpose. We have to be very careful about this.

Daniel Schmachtenberger: With regard to the new institutions we need, to what degree could existing institutions reform themselves, to what degree does it need to be new ones? It's kind of up to them. It's kind of up to the depth of realization of the need and the sincerity and then the coordination capacity of people in current institutions, how much role they could play. We can see the way that going into World War II, coming out of the Depression, the US up-regulated its coordination capacity so profoundly. So, could we have a Manhattan Project-like level organization? By organization, I mean the capacity to organize, not a singular thing, that was oriented to, how do we instantiate the next model of civilization? How do we instantiate the next model of social systems and social technologies? What is the future of education? What's the future of economics? What's the future of the fourth estate of law, et cetera, that fulfill the values that are meaningful and are anti-fragile in the presence of the current technologies and that can actually compete with the other applications of those technologies towards things that serve different values and/or aren't antifragile?

Daniel Schmachtenberger: I would love to see the US make that a central imperative, Manhattan Project level, to be able to do that. Not just how do we create a more powerful military, but how do we create a more powerful, a healthier, fundamentally, a healthier society that up-regulates and engages collective intelligence and its own problem solving and innovation better? I would like to see lots of countries do that. I think there are countries that did not yet transition to democracy, are interested in it, and can completely bypass the industrial-era democracies and go directly to better systems. I think networks of small countries, you see what Taiwan is doing, Estonia is trying to do some interesting things, I think networks of small countries could start sharing best practices and sharing resources so they don't all have to develop the stuff from scratch, which could start to lead to coalitions of countries like the EU saying, let's do some fundamentally better things.

Daniel Schmachtenberger: I think it will happen also not at the level of nation states, where decentralized groups, blockchain type groups say, "All right, let's really earnestly take on what these primary problems are and work on developing these solutions and these capacities." For the tech companies to do so would be very hard because while it could be profitable long-term, it would not be profit maximizing short term relative to the current thing they're doing. As we said, winning at the current game and building a new game are different things, and winning at a current game that's self-

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

terminating is a very short-sighted thing to want to keep doing. So if Facebook or Google or whatever were to cut its ad model, it would have a hard time being able to meet its fiduciary responsibility to shareholders a different way, but could it, in conjunction with a participatory government regulatory process that wanted to help change its fiduciary responsibility, where it became more of a social utility start to actually redirect its technology and redirect its decision-making process? Yeah, it could. That'd be super interesting.

Daniel Schmachtenberger: So I would like to see, as we mentioned earlier, I'd like to see the UN recognize that the level of progress that it has made at the sustainable development goals, nuclear deproliferation, and other types of international things like economic equality, globally, writ large, and preventing arms races and tragedy of the commons. That while it hasn't done nothing, what it's doing is not converging. It's not adequate, it's not converging on eventually solving the problem. It needs not just more of that approach, it needs a different approach. So to say, "Okay, well clearly we don't know how to facilitate coordination of global problems well enough, so let's have a superseding focus be innovation towards better methods of global coordination. That becomes our new number one goal, because we know we only get all the other goals if we get that."

Daniel Schmachtenberger: And you can see that during World War II, when we had to crack the Enigma Machine and figure out computation and whatever, we got Turing, we got Von Neumann, and we got all of the smartest people from countries all around the world engaged in solving those problems. I would like to see the US, the UN, I would like to see other countries, and I'd like to see private sector ,taking seriously the actual problem scape we have, and innovating not for just short-term advantage or narrow in-group advantage, but for long-term advantage of the whole. Since we have global effect, how do we build global coordination adequate to what is needed? To me that has to become the central zeitgeist and whatever groups figure out how to do it effectively will be the groups that can direct the future.

Tristan Harris: And I know that this is the work that you are working towards with the Consilience Project. Do you want to talk just about how you're working towards that with your work and how we're collaborating?

Daniel Schmachtenberger: Yeah, I mean, we're at the very, very beginning. The Consilience Project has a site up that is not even beta yet, just because we, in just starting, wanted to work on building stuff in association with thinking. But this talk is very central, this conversation you and I are having is very central to the aims of the Consilience Project, which is we're wanting to inspire, inform, and help direct a innovation zeitgeist, where the many different problems of the world start to get seen in terms of having interconnectivity and underlying drivers. And that the forcing function of the power of exponential tech is taken seriously, that says in order to become good stewards of that requires evolutions of both our social systems and our culture. The wisdom to be able to guide that power, recoupling, right, of wisdom and power, adequate to what is needed.

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

Daniel Schmachtenberger: So how do we innovate in culture the development of people, and how do we innovate in the social systems, the advancement of our coordination, both employing the exponential tech and being able to rightly guide it. So we have a really great team of people that are doing research and writing, basically, the types of things we're talking about here in more depth, explaining what is the role of the various social systems? Like what is the role of education to any society? Help understand fundamentally what that is, understand why there is a particularly higher educational threshold for open societies where people need to participate not just in the market, but in governance. Understand how that has been disrupted by the emerging tech and will be disrupted further by things like technological automation. And then envision what is the future of education adequate to an open society in a world that has the technology that's emerging?

Daniel Schmachtenberger: And we don't necessarily know what the answer is, but we know examples and we know criteria. So then it's like, innovate in this area and make sure you factor these criteria. And the same thing with the fourth estate, the same thing with law, the same thing with economics. So the goal is not how do we take some small group of people to build the future? It's how do we help get what the criteria of a viable future must be? And if people disagree, awesome. Publicly disagree and have the conversation now, but if we get to put out those design constraints, someone says, "No, we think it's other ones," at least now the center of culture starts to be thinking about the most pressing issues in fundamental ways, and how to think about them appropriately and how to approach them appropriately. So fundamentally our goal is supporting an increased cultural understanding of the nature of the problems that we face, a clearer understanding, rather than just there's lots of problems and it's overwhelming and it's a bummer

Daniel Schmachtenberger: and so either some very narrow action on some very narrow part of it makes sense, which is most of activism, or just nihilism. We want to be able to say, "Actually, because there are underlying drivers, there is actually a possibility to resolve these things." It does require the fullness of our capacity applied to it, and with the fullness of our capacity. So it's not a given, but with the fullness of our capacity applied to it, there is actually a path forward. So we're writing these papers that basically would be kind of like a meta-curriculum for people who want to be engaged in designing the future. And some of them have to do with current public culture and how to be able to change patterns of public culture that lead to better conversation, better sense-making, better meaning-making and choice-making, so there is an on-ramp into higher quality conversations, meaning higher quality process of conversation.

Daniel Schmachtenberger: And then some of them are things like what are the design criteria of the future social systems and how could we build those things? Then, not everybody will read those. Some people who have the ability to help start building them will. But we hope that other people will take that and translate it on podcasts and into animations and in whatever other forms of media so that those topics start to become increasingly present in people's awareness. Then of course the next part is, what groups start emerging

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

wanting to address those and what can we do to help facilitate good solutions in those groups?

Daniel Schmachtenberger: And this is where you and I... I've learned a lot from you about the social media issues in particular and how central they are to the breakdown of sense-making, because obviously without good shared sense-making, there is no possibility for emergent order. You either just get chaos or you have to have imposed order. If you want emergent order, that means emergent good choice-making, that means emergent good sense-making. So we've learned a lot and discussed these things for a long time, and obviously also not just you and I. There's a whole network of people that we're connected to that have been thinking deeply about these things and that we continue to try to think about what adequate solutions could look like.

Daniel Schmachtenberger: I think what CHT did with The Social Dilemma took one really critical part of this meta-crisis into popular attention, maybe in a more powerful way than I have seen done otherwise, because as big a deal as getting climate change in public attention is, it's not clear that climate change is something that is driving the underlying basis of all the problems, but a breakdown in sense-making, and a control of patterns of human behavior that kind of downgrade people like, "Oh, wow, that really does make all these other things worse." So I see that as a very powerful and personal on-ramp for those who are interested to be able to come into this deeper conversation.

Daniel Schmachtenberger: And some of them it'll simply help them be like, "Okay, now I know what I was intuitively feeling. Somebody's put it into words and I at least feel more oriented." And that's the extent, because they don't necessarily have the ability to build new blockchain systems or whatever it is, and they should be doing the nursing or education or whatever really other important social value they're doing. Some people will be able to say, "This actually really resonates. I can translate this to other audiences, and get more people engaged." And some people say, "I can actually start innovating and working with this stuff." And all of those are good.

Tristan Harris: Yeah, I agree, and I think what we've essentially been outlining here, and you hit it at the end, is going back to the Charles Kettering quote, which I learned from you, and I've learned so many things from you over the years, which is that a problem not fully understood is unsolvable, and a problem that is fully understood is half solved, and I just want to maybe leave our listeners with that, which is, I think people can look at the long litany of problems and feel overwhelmed or get to despair in a hurry, I think is your phrase for it, and I think that when you understand the core generator functions for what is driving so many of these problems to happen simultaneously, there's a different and more empowering relationship to that, and you've actually offered a vision for how technology can be consciously employed, these new technologies can be consciously employed, in ways that should feel inspiring and exciting.

## Center for Humane Technology | Your Undivided Attention Podcast

### Episode 36: UNEDITED A Problem Well-Stated is Half Solved

Tristan Harris: I mean, I want that transparent blockchain on a budget for every country in the world, and we can see examples like Estonia and Taiwan moving in this direction already. And we can see Taiwan building some of the technologies you mentioned to identify propositions of shared values between citizens who want to vote collectively on something that previously would have driven up more polarization. So we're seeing this thing emerging, and I think what we need is to sort of have this be seen as a necessary upgrade to, again, I think we need to see this as not just an upgrade, but the kind of cultural enlightenment that you speak of that so many different actors are in a sense already working on. We used to have this phrase that everyone is on the same team, they just don't know it yet, and once you understand the, I think, degree to which we are in trouble if we do not get our heads around this and identify the kind of core generator functions that we need to be addressing, once we all see that ...

Tristan Harris: I'll just speak to my own experience. When I first encountered your work, and I encountered the kind of core drivers that drive so much of the danger that we are headed towards, I immediately, I was kind of already in this direction already, but I reoriented my whole life to say, "How do we be in service of this not happening and of creating a better world that actually meets and addresses these problems?" And I know so many other people whose work and whose lives and whose daily missions and purpose have been redirected by, I think, hearing some of the core frames that you offer, and who I hope and who are... Many of whom are already working on active projects to deal with this and those who are not are supporting in other ways, and I just hope that our audience takes this as an inspiration for how can we in the face of stark and difficult realities, as part of this process, gain the kind of cultural strength to face these things head on and to orient our lives accordingly.

Tristan Harris: Because I have, during periods of time, hit probably low-grade despair myself. I actually feel more inspired than ever the amount of things and the number of people who are waking up to these challenges. And I'll just say that, I think when you face these challenges alone and you feel like you're the only one seeing them or you have a weird feeling in your stomach, it can feel debilitating, and when you realize the number of people who are also putting their heads up to say, "How can we change this," that's what feels hopeful, and that's where I derive my optimism. So Daniel, thank you so much for coming on. It's an honor to have you. Your work has touched the lives and work of so many people who may not always say so publicly, but I know that you had also a huge hand in inspiring some of the themes that emerged in The Social Dilemma, which has impacted so many people as well. So thank you so much.

Daniel Schmachtenberger: Really wonderful that we could have this conversation. Thanks, Tristan.

Tristan Harris: Absolutely.