URL: https://ecorner.stanford.edu/?post_type=snippet&p=55046

Tristan Harris, a tech entrepreneur turned design ethicists, offers a framework for better aligning today's systems with human values. Using artificial intelligence as an example, Harris calls for a more symbiotic relationship with AI, instead of the current state, where it is designed to extract our attention. He likens the relationship to that of depleting Earth's natural resources, and calls for a more sustainable model based on renewal and replenishment.

Transcript

- And we talk about these future scenarios where if you have runaway AI, there's always this discussion of like, "Well, let's put it inside of an air gapped computer "and it won't like, get out," right? Or, "We can always pull the plug.. "Let's just shut the thing down.. "We can always pull the plug, "and it's not gonna be a big problem "because we can always just turn the computer off "if we have this runaway AI in the future." But the problem is, if you think about what that would be today, that would basically be like turning the lights off at Facebook.. Now, it's not as if, from an existentialism perspective, as a human being, you can't like, get out of your chair and walk out of the room in space and time and tell the board of directors, you know, "We're gonna shut this thing off." You could physically do that, but there'd be consequences, and we can't turn off the system that we've created.. So, the only way to solve a system that's run away pursuing its own goals is not to manage your relationship to it.. I'm saying this 'cause I wanna contextualize.. A lot of people think of, if you've read any of the work that I've done, that it's about we oughta better manage our relationship to our phones.. Let's be more mindful with our phones.. That's like saying when chess is kicking your butt, you know, "Let's manage our relationship to the chess "that's just totally overpowering where we are." So the only way to solve this problem is actually put the AI on the same side of the table as us.. We have to be really honest about what the lines of power are..

So right now, the thing is pointed at us.. It's extractive.. It basically says, "I need to extract "as much attention out of you as possible," and the classic line is, "It's a race "to the bottom of the brain stem," so, it's not enough that I offer you the product, I have to reach deeper down in the brain stem and create an unconscious habit so now you actually pull for the phone more often.. That's not enough.. I have to reach even deeper down into the brain stem and own your social psychology, so the way that Snapchat basically owns people's social relationships.. That's not enough.. I have to reach deeper down the brain stem and get to your self-worth and control your sense of how often people like your self-worth.. So I told you I'd go dark, I apologize.. So we have this extractive economy, and if you'll notice, it doesn't look that different from how the broader economy is structured when it's extractive.. It's the same as the environment..

We have a runaway system that makes more money the more you extract without putting something back in balance, so the fundamental situation here is we need to find a more ergonomic relationship with the boundaries of our architecture in the case of how it's extracting us, and if you don't care about us, let's talk about our kids because this thing is just eating kids alive for breakfast.. So, we can do that, but I think that what this is about is just like the point where we were where we realized, you know, after we extracted coal and we created this incredible economic prosperity after generating energy with coal, but we also polluted the external environment, and no one really saw that at the time.. Good-intentioned people created this bad stuff and we just realized it later.. And we're at that point where we're just realizing this extractive attention economy where we're capturing people's attention is basically polluting the external environment.. It's also polluting the inner environment, and now, we need to invent solar and green basically solutions, like things that are regenerative or ergonomic or replenishing to the human being, so you can make all the same metaphors in the environment, and that's gonna mean that things don't grow the same way.. One of the reasons why advertising is such a great business model is because it grows, you know, incredibly fast and is super profitable.. I mean, who wants Facebook to be regulated when the stock price is like, through the roof right now? But also, who wants the changes that we're gonna need to make for climate change to be abated, given that we're gonna have to make huge sacrifices, and so the conversation actually comes back to values.. It comes back to, "Am I willing "to value something else besides money? "Am I willing to value something else besides status?" Now, this is

an open question when it comes to corporations who basically, they don't have control or a choice.. We created these robots called corporations that are just maximizing profit, and they have to.. There's no like, way they can't do that..

So we have to acknowledge when that misalignment exists, and it's the same for our own lives.. You know, just because I can start a company in my friend's, you know, I have friends personally who, you know, sold a company for a billion dollars.. Now, that's very seductive.. We can be pulled.. It's a magic trick.. We can be pulled by basically other people's success or incredible amounts of money, right, but the conversation always comes back to, what do we care about? What is this for? What is the problem that this technology is the solution to? Why am I doing this? And if we ask those questions, that's really the essence of ethics...