Center for Humane Technology | *Your Undivided Attention* Podcast

Can We Govern AI?

Tristan Harris: How can you govern something that's uncontrollable or acts in ways you can never predict? A few weeks ago, I joined thousands of people in the technology sector who signed a letter asking companies to pause on releasing any more AI large language models, and to give us enough time to respond to what's out there so that we can move at the speed of getting this right.

I'm Tristan Harris, and this is *Your Undivided Attention*. Now, a pause or slowdown is just one thing that we can do when it comes to AI, but another is government regulation. And lately we at the Center for Humane Technology and others have been asking, what would that even look like? One potential model is the European Union, which enacted a whole raft of tech rules starting about six years ago, and they have several more in the pipeline, including one called the AI Act. Now some of them have real teeth. For example, Meta or Facebook is actually considering banning political ads in Europe, specifically because it would be too difficult to comply with the Digital Services Act campaign advertising regulations.

Now, I'll admit that I'm a bit skeptical of these things right now. It's well known that the US lags behind the EU when it comes to regulation, let alone the fact that AI presents such a new kind of risk with speed and complexity that's so hard to imagine even existing laws catching up to, let alone getting ahead of them. Now, my guest today feels completely the opposite. Marietje Schaake was at the center of developing a framework on tech regulations in her former role at the European Parliament, and she argues that there's a way for 21st century regulation to stand up to the threat of runaway AI.

Welcome to *Your Undivided Attention*. I am so glad to have here with me today, Marietje Schaake, who's international policy director at the Stanford University Cyber Policy Center and an international policy fellow at Stanford's Institute for Human-Centered AI. And she's also a former member of the European Parliament for the Dutch Liberal Democratic Party where she focused on trade, foreign affairs and technology policy. And I remember when we first started actually working on social media in 2017, we had a meeting at Common Sense Media's office in San Francisco, and we were just starting to talk about, gosh, how would you even regulate social media? And now here we are, whatever it is, six years later, and we have many more problems on our hands, but welcome to *Your Undivided Attention*.

Marietje Schaak...: Thank you, Tristan. It's really great to be here.

Tristan Harris: So let's just get started with introducing you a little bit to our listeners. Who are you and what was your role in creating some of the tech regulations in the EU?

Marietje Schaak...: Yeah, so I am now working on tech policy only at Stanford University, but I come from the practical lived experience of making laws and making policy as an elected parliamentarian for the period of a decade. And while I was in the

European Parliament, we adopted a whole bunch of laws because in Europe the thinking about regulation is actually far more advanced than it is in the United States. And also not just the thinking, but also the doing. And so, for me, the need to put in place guardrails, checks and balances, oversight mechanisms is normal and we should also normalize it. It is not an attack on tech companies or Silicon Valley that Europeans want to do this. It is actually a very normal response to the growth of an industry and in particular the urgent need now to mitigate all the harms that I know you've worked on so intensively.

Tristan Harris: Yeah. So why don't we take a step back and ask what even is regulation? Why do we need guardrails on this?

Marietje Schaak...: Regulations are essentially rules that everybody should adhere to, and I think it's really important to keep that in mind that laws are not only there to protect people from the outsized power of companies, tech companies, but also to protect people from the outsized power of government. And in the discussion about tech policy that is often lost, it often seems like the governments or the lawmakers, Congress in Washington is just out there to make life miserable for companies, to take away the fun services like TikTok, a very current discussion that we've had where you see all these content makers saying, "Don't take away our business, don't take away the fun of our teenagers." But obviously just showing the entertainment value or the market value does no justice to the harms that you talked about.

So I think of regulation as a level playing field, the same rules that apply to everyone and that create a bottom line, the lowest sort of necessary safeguards for public health, public safety, wellbeing of people, the protection of children, the protection of the common good. So I actually think regulation, if done well, is great. It is what guarantees that we live in freedom and that also the rights of minorities, for example, are respected.

Now taking that to AI, what kind of regulations might we need to deal with this rapidly developing new class of technologies? I think there are a couple of fundamental challenges to navigate that make AI different than other technologies, but also other products and services that have been regulated before. One is the information about the use of the technologies, but also the datasets going in to change them is not accessible to lawmakers, to journalists, to you and I. It is in proprietary hands. These companies guard the secrets to their algorithmic settings with their life.

The second thing is that with the constant new iterations and the very personalized experiences that people have, the product or service is fluid. You can't hold it, you can't pinpoint it, hold it down. It is different for you than it is for me. It is different today than it was last week. And so imagine being a regulator that is supposed to establish whether illegal discrimination has taken place or whether consumer rights have been respected, where do you begin?

And so with the combination of a lack of access to information and the fluidity of the service and the product, that makes it very hard to regulate.

So maybe I'll leave it there for now to give you a sketch of what I think makes AI and AI regulation specific and particularly challenging compared to, let's say, pharmaceutical regulation.

Tristan Harris: Yeah, I think it is helpful to establish a baseline of other kinds of regulations that are much more straightforward or easy to do. We think about pharmaceuticals, which also have unpredictable effects on the body or interaction effects with other pharmaceuticals. And so there is sort of an interesting parallel there where you release social media into the world, maybe it works well for an individual user and there's no obvious harms. Those don't emerge as discrete harms like Twitter caused this prick of blood to emerge from my body where something actually went wrong or a drug that has an adverse side effect where I get a stomach ache or something like that. So I think it might be helpful maybe to set some ground on what makes regulating social media or AI or just runaway technology in general different than previous classes of, let's say, airplanes or pharmaceuticals or food.

Marietje Schaak...: Let me start by where they are the same. I think nobody would ever say that regulation is easy. So even if we think that regulating AI and other technologies is hard, think about chemicals, think about financial services, think about food, the enormous complexity, wide variety, constant innovations that happen in those sectors too. I mean, there's constantly new combination of chemicals, of foods, of financial services. So we should not be discouraged is what I'm trying to say by the fact that a problem is complex.

Tristan Harris: Great point.

Marietje Schaak...: We should trust that we can really make it work. And it's also high time that we make it work for a number of these technologies because it is entirely normal that there are rules to be safe and to, for example, also have a place to go when you've been wronged. Let's imagine you've been poisoned by the use of a medication. Well, then you want to go somewhere and not just be left on your own with all the harms that it's done to you. So in that sense, I think we need to normalize tech regulation and not see it as an exceptional set of problems that cannot be solved. It will just require unique steps just like the chemicals and the pharma and the food have required unique steps.

What does make it somewhat different is the global nature of companies, the fact that they may operate from one jurisdiction, but they reach consumers, users, internet users, citizens completely in a different context on the other side of the world where that different context creates different circumstances and can make people vulnerable, can lead to all kinds of new problems.

Tristan Harris: So I think you're bringing up a great set of points here on what's similar to existing issues that we've faced before, chemicals, regulating lead, regulating DDT, regulating food and drugs, regulating airplanes, nuclear power plants. These are all areas where there's complexity, where the people who are regulating don't have the same knowledge as the amount of complexity that's inside of chemistry or pharmaceuticals or what goes into a nuclear power plant. So we've dealt with that problem before. That's the lack of knowledge issue of regulators.

Then there's transparency. Can there be some notion of what we know about the industry and its practices and is it a black box and are we allowed to investigate, or is there no way to investigate Twitter's algorithm or Facebook's algorithm or the safety practices of a nuclear power plant or something like that?

Then there's a unique challenge which is the shapeshifting challenge. Now you said, food companies or drug companies may update the formulas of what goes into a Cheez-It, five years ago may be different than what goes into a Cheez-It today and whether they use Roundup or other kinds of nefarious ingredients or something like that. But technology, as you said, sort of shapeshifts much faster.

And one of the things I do want to outline for listeners that I think does present a unique challenge with this class of digital technology and whether it's AI or social media, is this notion of a complexity gap, that the pace, speed and scale and complexity of technology is updating much faster. If you imagine graphing, it would sort of move up at an exponential compared to culture and governance, meaning how much does the culture understand the new technology and it's monitoring that complexity and also how much are the institutions and the regulators or the governance process able to understand that new complexity. And anywhere in that gap between the actual complexity of technology's impact on society and our understanding or our governance of it, anywhere in that gap is where externalities, risk, pollution add up on the balance sheet of society and aggregate into a kind of net fragility or existential risk that kind of destabilizes some of society.

So of all the kinds of regulations that the EU has looked at over the last 10 years, what are the most important ones to deal with the challenges that we are talking about facing?

Marietje Schaak...: So let me start by saying that the bulk of regulations that are going to be very significant in dealing with AI and other emerging technologies are still in the pipeline. So I can point to some of the laws that have been adopted and implemented and enforced over the past decade, but I think the next decade will be crucial to look at.

There are a variety of laws that matter a great deal right now. Think about an old but very fundamental framework that is quite similar in the US and the EU, and that is that of antitrust and competition law. It is intended to protect consumers and to avoid the abuse of power by companies in the market. It doesn't deal with all of society, which is why other laws are needed, but abuse of market power to squeeze out competitors, to buy out startups, to take away the oxygen out of the market and innovation is a problem that a lot of big tech companies have been guilty of and that they are now facing scrutiny for on the basis of these laws that are about a century old.

Then in the EU, there have been important steps made in terms of data protection, which I think some listeners may be familiar with the GDPR, the General Data Protection Regulation, which has been adopted in 2018 and is one of those examples that horizontally applies. So it is supposed to protect people against the abuse of their data and the mishandling of their data by tech companies, but also by governments.

Tristan Harris: And then there's the Digital Services Act and Digital Markets Act. Can you explain those to our listeners?

Marietje Schaak...: Yes. So there's a twin set of regulations that each address related but different aspects of the business models of tech companies. The Digital Services Act deals with content. So, for example, hate speech or disinformation or harassment. The Digital Services Act wants to create mechanisms that puts in clear words what the responsibilities of tech companies are in terms of dealing with this content, removing this content and what these processes look like. And on top of that, there are transparency requirements about the algorithmic settings of companies, which I think make it really interesting and requirements to provide data so that independent researchers like academics here at Stanford can actually look at them. And those are all opportunities that Americans don't have at the moment.

So the Digital Markets Act tries to make the responsibilities of particularly gatekeeper companies much more proactive. So it spells out how they should behave, not just after they've violated the law, but it creates a clarity on their expectations so that there is more fairness in the digital market. That's the goal. And so you can think of the Digital Markets Act and the Digital Services Act as dealing with speech and content issues, trust and disinformation issues, and the market power issues that these big companies have. And so it really does begin to chip away at their business models. It won't be a sort of end all and be all, but it really comes at the problem of the outsized power of big tech companies from two different and important sides of speech and content and of competition and market power.

Now enforcement is the crucial part. You know how new laws are often announced with great press releases about how everything is going to change, but really the success of a new law, particularly the General Data Protection Regulation, but also in the future, the AI Act, which is one of those laws in the pipeline, will really depend on the rigor and the effectiveness with which enforcement happens. And here we see an imbalance. The budgets available for enforcers are very, very small. The ability to hire the top-notch talents is very, very hard as a result.

And so I'll give you one example that really made me want to pull my hair out. The Dutch Data Protection Regulation recently added €1 million, which is slightly more than $1 million to its budgets to deal with algorithmic oversight. A million. Okay, well, good luck. So here's your extra million vis-a-vis billion-dollar companies like YouTube, Amazon, Twitter, Meta. What are you realistically hoping that that is going to do?

And so, one way to think about how we might use the good of regulation but improve it in light of the need for this 21st century governance that you speak about is to sort of turn the model on its head. And let me explain what I mean by that. Typically, in regulations there is quite specific framing of what a law entails. So what behavior is allowed and is not allowed, and then there is relatively little investment in the enforcements.

Now, I think if we try to articulate laws and regulations more in the form of principles that need to be safeguarded and less specifically about the details of technologies, but then double down and not even double down, but tenfold, 20-fold, 30-fold. Empower the enforcers with stronger mandates but also stronger abilities to enforce, the ability to hire, the ability to probe, the ability to sanction will really give them the power that is needed to, for example, ask for information that is typically held in secret by tech companies to really understand details of how algorithmic settings can impact society so that when there is access to information, people know what to do with it, which requires specific expertise and knowledge.

And then give them the power to really hand down sanctions that bite. And that's the last thing I'll say about efforts in Europe with the Digital Services Act and Digital Markets Act that are also effectively still having to prove their effectiveness. But still the thought there is that the sanctions should really hurt because in the past, also with antitrust, even if a fine of, let's say, $3 billion was handed to a tech company, that may well have just been written off as the cost of doing business. Twenty years ago, $3 billion would've been an astronomical amount of money to be fined for any commercial company. Right now, if you have profits and turnovers of what, like $100 or $200 billion, what pain is $3 billion going to cause?

And so what they're doing now is to have a percentage of profits as the sanction. So the sanctions are now proportionate to the market power. And I think that that is another way to think about it and to have the stick that really can impact the bottom line of the companies. And so for democratic governments to sort of meet that power, it is important that they collaborate. Because in the EU we often see that fragmentation between different governments or an unpreparedness is used by companies to try to sow division between the different governments, and that's all entirely predictable.

Tristan Harris: Could you give an example of how companies' lobbyists will use to create division between governments?

Marietje Schaak...: Well, for example, when you look at the tensions around the use of network technology companies from China, like Huawei or ZTE, in Europe, there was an emerging discussion about whether these companies were safe and in the meantime, the lobbying and the pitching for contracts and the ongoing deal making was going on. And so you will see that whenever there's a difference in how rules are made or interpreted between different governments, we could also look at the US and the EU right now, companies benefit because the scale of the level playing field where companies have to adhere to the same rules and the leverage that governments can have is simply minimized.

You can see that with everything. If you are small groups protesting something, you make less of an impression than if you are a big group. And I think at this point in time, the combined impact on society of some of these emerging technologies who are already present, powerful, entrenched business models require a meeting of minds between like-minded governments to really work together because otherwise, because these are globally operating companies, they can easily swim through the holes in the net and they will. So that's what I'm trying to say, that if you're recognizing that scale and power is a big factor, you want to have scale and power to meet the problems.

Tristan Harris: Symmetry of power is the principle.

Marietje Schaak...: Yes, you're so right. And so the EU typically will leverage its shared market. It's one of the biggest markets in the world. That's the whole idea behind the EU as well, to have a single market where the same rules apply, and that's not only to facilitate commerce, but also to make sure that it adheres to the same rules, toy safety for all toys coming into the EU. And so the same should apply for tech. And what you see in practice, and that's I think where it gets interesting is that, for example, in the case of the General Data Protection Regulation, there were companies like Microsoft who just realized if this is the most strict set of rules that apply globally, we might as well adhere to them everywhere we operate because then we're never wrong.

And so the sort of extraterritorial effect or the first mover advantage, if you want to think about it that way, can sometimes have a global impact. And so the EU is aware of this and is, I think hoping to replicate that with Digital Services Act, Digital Markets Act and the AI Act in that it's hoping that new high standards that the EU is first to adopt will actually incentivize companies to adhere to them worldwide-

Tristan Harris: Over the world.

Marietje Schaak...: Yeah, which should also make it easier for businesses to operate in the EU and then maybe would also make it easier to discuss shared engagement by, let's say, the US government and the European Union to actually set up these rules. There's a dialogue going on called the Trade and Technology Council where topics like AI but also quantum and cybersecurity and internet of things and all kinds of emerging technologies are on the table to see if they can align what they do in terms of-

Tristan Harris: Align the US and the EU sort of policies so that there's an even bigger net that becomes the standard one.

Marietje Schaak...: Yeah.

Tristan Harris: Yeah, I totally hear you. There's a joke, I think, about how the US brings the software, China brings the hardware, and EU brings the regulation or the laws to bind them. I'm not sure if that's really a great joke or not. But one of the notions that you're talking about is, I'm thinking about the Star Wars metaphor of "Help me Obi-Wan Kenobi. You're our only hope." When I go to Washington DC and we talk to government policy makers about the need to deal with AI issues or social media, oftentimes they just look at you so depressed and like you're going to have to go to the EU or California state to get something to happen because here we're so divided and there are natural and organic reasons for that division, but as you said, companies will also weaponize division.

And you'll see this over and over again that what happened with social media was instead of a debate about the engagement for a profit business model, which is the root of the issue and the concentration of power around that, the social media companies weaponized one narrative so that the Republicans in the US became concerned with censorship and free speech, and the left became concerned with misinformation. If I forecast forward to the risks with the AI conversation, I think the biggest risk for division is around making it about what kind of bias does the AI have, because that will create, at least in the US a division between the left and the right. So yeah, I really appreciate laying all this out because I think it speaks to how do we get to a world where there is that symmetry of power?

Marietje Schaak...: I think we should not expect that a variety of voices will suddenly all come together in our democratic systems. There will always be differences in priorities, differences in understanding, and a need to achieve compromise. And that in the US is very difficult right now. It's not easy in Europe either, but indeed that's necessary and it would help to explain the problem in multiple ways. So, for example, why not dealing with harms impacts children but also impacts the environment or fairness, a fair opportunity to do business and so on.

So it's important to imagine how different players in the political spectrum might relate to a topic in order to bring them on board. And I think in the US, something interesting is happening right now. Indeed, the EU was long the sort of sole regulator of big democratic countries to deal with the outsized power of tech companies and to make sure that rights were protected, that fair competition was there, that there was consideration of security. But the US is really catching up in a maybe unexpected way, or at least in a way that I am only now beginning to appreciate, which is the lens of national security is what is bringing Republicans and Democrats together.

And as a result, you see that Joe Biden and his administration are adopting executive order after executive order to reign in the tech sector. There was the announcement of a ban of the use of commercial spyware by the US government. There have been export controls announced. There is the whole saga around TikTok and what is and is not safe, the decoupling with Chinese technologies because it's not trusted and seen as a strategic tool in the hands of the Communist Party. So I am seeing a huge catching up in the United States anchored in very different values and motivations than what I see in the EU.

And on the one hand, I'm happy that there is an awakening in the United States that the outsized power of the market is dangerous also for national security. What I worry about is that the whole dynamic of civil rights protections, protecting the public interest, making sure that there is transparency and accountability just because those are principles of a rule of law-based society are kind of brushed aside to really push this more opportune, this more popular, this more bipartisan narrative of national security.

And so I think it's a delicate balance to strike where it's important that national security does not become the justification to also sacrifice civil rights as it has been. After 9/11, mass surveillance was justified in the interest of countering terrorism, discrimination against minorities became normalized. So this is really a moment to pay attention to what is happening in the United States in terms of tech regulation. Even if Congress may be dysfunctional, it doesn't mean that nothing is happening.

Tristan Harris: And this does point to models like Audrey Tang's digital democracy as why something like her digital 21st century governance, a system in which citizens are actually entering the statements about what kind of way do they want to relate to social media, where they get the benefits of small and medium-sized enterprises being able to efficiently reach people, but also not having children's harms and not having polarization and not having overzealous censorship of ideas. There is a unity of those views, and we need digital systems that actually help us find those things. And what we're needing is we need examples of governance that also uses technology in a way that allows it to evolve and move at the pace and speed of the issues of technology. And I think Audrey Tang's digital democracy is the best optimistic and exciting example of that that I've seen.

Marietje Schaak...: Yeah, she's absolutely very inspiring, but I also see a sort of delicate balance there, where sometimes the narrative of service delivery by governments is used to push the quicker adoption of technologies by big tech and to entrench governments themselves. So there's this whole narrative that basically governments are completely dysfunctional because they are not with the 21st century and so on. I think there are a variety of ways in which this moment can be met with making laws and hearing people's voices and making sure that there is inclusive representation and a consideration of which communities are impacted the most by certain technologies, but it shouldn't distract from the need to really have a principled-based approach to achieving transparency, access to information, oversight, rights protections, and resilience of systems too, because we haven't touched much on cybersecurity and the vulnerability of the technologies that are sold as the best solutions.

Tristan Harris: So breaking in here, what other kinds of regulation might we need and what are some of the principles that we didn't cover in this interview that we might want to talk about? Well, here's some of the frames that we've been hearing from people in the space.

People right now are talking about regulating GPUs or graphics processing units, which are the fundamentals for how these AI companies train their new large language models. Should we have auditing regimes that know if you're going to do a large training run? Should that be a licensed procedure just like you have to apply for a license to become a doctor or a lawyer? Should you have to have a license to run an autonomous language model on a cloud provider? Should we have KYC or Know Your Customer laws just like we have for banks? If you create a bank account, you have to put in your address and social security number so the bank knows who you are.

Similarly, should we have know your customer laws for cloud providers where they have to know who's using those cloud providers to train large language models so that we can keep track of who's doing this kind of work? Should we have a liability regime? Should the AI companies who train models be liable for the kind of downstream damage that they do? Should cloud providers that hosted a model that was starting to wreak havoc that caused damage down the line, should that cloud provider have some kind of liability for hosting?

In general, the principle that we have to deal with is power has to be matched with responsibility. Let's take a second and talk about how this works in the medical sphere. We know the kinds of powers that a doctor has to alter the physiology and health and wellbeing of a person. And because we recognize those are special powers, we developed a hippocratic oath and a white lab coat and a ceremony and a ritual for imbuing a sense of responsibility into the doctors that put on that white lab coat for the first time.

But AI engineers who could be 18 years old are just people who are hacking around on GitHub and downloading code around on their laptop. They don't have an embodied sense of responsibility and moreover, we don't know all the things that AIs can do. What does it mean to be responsible when you're creating something that has more capabilities than you know what you would need to be responsible for? What would constitute safety? We know that for cars and airplanes that there's certain things that we found out have to do with what makes a car safe, what makes an airplane safe, and when there's a car crash or an airplane crash, we iterate and improve those standards of safety so we can make cars and airplanes increasingly safe.

But can we create safety standards after an AI crash? Let's say we're talking about AI taking control and running away and becoming like a virus that copies its code around the internet and runs itself on more and more cloud providers and is able to keep doing that in an uncontrolled way. Do we get a second chance if we get that wrong? Artificial intelligence poses brand new questions for what would constitute safety in the digital world, much more so than cars or airplanes, and that's why it's such a unique and profound regulation challenge. And now, back to the interview.

Would you like to Steel Man both the perspective of regulation not being up to the task and regulation being up to the task that we can build on this foundation and trust it for the next 10 years?

Marietje Schaak...: So when we think about regulation, I think it's important not to see it in black and white terms, which has been the case for a long time. Questions have been prompted to CEOs where they're asked, are you in favor or against regulation? Are you in favor or against rules for technologies? And obviously the question is it depends on which kind of rules. And when we look at the General Data Protection Regulation, for example, or the laws around cookies in the EU, there are a lot of lessons learned about how things could have been done better. For example, the GDPR doesn't really take into consideration AI as such. There's an enforcement problem that I pointed to earlier. But what I believe is a little bit unfair in the assessment of how well regulations work, is that it's very easy to say where they're failing, whereas we don't talk as much about where the status quo of nonregulation is failing. And I think we need to do that too.

Now with regard to the new laws that are almost ready to be tested for their merits in the EU, the Digital Markets Act, which spells out more clearly what the responsibility of tech companies is when it comes to fair competition, with regards to the Digital Services Act that talks about content moderation and the responsibility companies have, with regard to the AI Act that tries to mitigate risk in the way which AI is used in society. For example, if the impact is a loss of liberty or a loss of employment or a loss of access to social services or education, that is deemed a high risk application of AI.

In other words, a risk-based approach to the use of AI is something that I think is a good step. It may not be covering everything because we're learning new things about AI every day, and we've already seen major breakthroughs in large language models and the way that they are being used since the AI Act has been initiated. So we need to have flexibility built into regulations in order for them to be able to expand, to cover new threats and challenges coming from new technologies.

Tristan Harris: That's a perfect example. I mean, I'll admit, I think that's my deep fear is that just speaking really openly, like the advancement of large language model AIs, which we did a big podcast episode on called the AI Dilemma and why the double exponential curve of improvements and the unknown capabilities that pop out of these large language models, the more you feed them data, the more unexpected capabilities that they have to write code, to be jailbroken, to speak other languages, to hide information from their own users, to deceive their trainers. There's lots of things that they can suddenly do. And those capacities, as you said, were not envisioned as part of the EU AI Act because that started to get worked on before these capacities and these properties of large language models were known. So how do we actually get ahead of these risks?

How do you relate to that question? Because honestly, I'm really, really, really concerned that we don't have levers that are currently available to us that can create the guardrails that allow us to get this right, which is why you saw things like this big letter that I signed along with Elon Musk and Steve Wozniak, the co-founder of Apple, that we just need to pause these large language model experiments to give us enough time to respond. How do you think about that?

Marietje Schaak...: Well, so I think the pause is a nice idea, but I don't think it's going to happen because the race that you spoke about earlier has already taken off. And even if some companies may opt to pause, which I doubt will happen, then others may still want to race ahead. And so it won't create a comprehensive pause. I want to look more to the seeds of opportunity that exist in regulation, and again, I want listeners to feel empowered that we can change the status quo.

So in the AI Act, there is also an AI board foreseen where a group of experts delegated by member states and the European Commission would have the opportunity to look at new applications of AI and assess their risk. So that's a seed of opportunity, right? It will depend on how this board will consider its own mandate, whether it can agree that, for example, these existential risks of large language models, or rather maybe less existential, but still very practical risks of being wrong, of being harmful, being discriminatory, being deceiving, being biased, are all happening, and that that might be a reason to deal with them a certain way.

Similarly, we've seen that in Italy, the data protection regulation has challenged the use of data feeding into these large language models on the basis of the General Data Protection Regulation. That is a very important kernel of opportunity. If it turns out that the data that has been scraped from left, right, and center, I mean these companies are hoarding data in order to train their models. If it turns out that they have done so in an illegitimate way according to the EU law, that is a huge game changer because that will really allow accountability.

Lastly, what I'll add to this to come back to your notion of the AI moves so fast, the regulation moves so slow, almost every law or regulation has come in response to a wrong or a harm in society. It is impossible in any case to anticipate the future, even if people try. There has been a lot of scenario planning in anticipation for pandemics, and yet COVID-19 came as a surprise to many governments and left basically the world unprepared. Now people are preparing for the next pandemic, but undoubtedly the problem will come from elsewhere.

So what I'm trying to say is the fact that regulation follows new realities that new technologies create is normal, it is not going to change because you cannot regulate for the unknown. What I believe you need to do is to empower people to look for emerging risks and challenges as the technology goes along. And that means not trying to list every possible risk around AI or other technologies today, and to hope that that will sustain you for the next decade, but rather to say what we know is that there's a lot that we don't know, and we're going to empower people to look for effects of the new technologies, how they impact existing rights, how they create new risks, new realities, new challenges, new harms to the common good, to the environment, to young people, to education, to democracy, what have you, for these mandated experts to then come forward with their analysis and the enforcement agencies to be able to intervene.

And I think we can do it as long as there's a political will. And before we sort of conclude, the one positive thing of the enormous impact and risk of large language models and other kinds of AI applications that we're seeing now may well be that the awakening and the wake-up calls that took a long time for social media to hit society will come faster because it is so clear how disruptive these systems are, that it may actually lead to political will to intervene quicker. Now, the only challenge, and it's a big one before us, is to make sure that the political interventions are hitting the right spot and that it's not a kind of do something reflects that will lead to good intentions, but bad outcomes.

Tristan Harris: Yeah, that's totally been our theory of change that we have to leverage the now recognized cultural understanding that had we gotten ahead of social media becoming entangled irreversibly into our society, that it would've been so much easier to regulate social media before it got entangled with elections and politics and journalism, and become the basis of how small and medium-sized businesses reach end users. Had we gotten ahead of that, we could have actually regulated social media because there would be many fewer vested interests competing. And with AI, we're really hoping that people can see that and that we can get ahead of it and regulate it now.

Marietje Schaak...: Yes.

Tristan Harris: I know we're out of time. I'm so grateful for you taking the time to be here with us on Your Undivided Attention and to walk through in the "Help me Obi-Wan Kenobi." The EU regulation might be our only hope for it right now because the US has not been able to act until recently on issues of national security. So thank you so much for coming.

Marietje Schaak...: You're welcome.

Tristan Harris: Marietje Schaake is the international policy director at Stanford's Cyber Policy Center and a fellow at Stanford's Institute for Human-Centered AI. And before that, she was a member of the European Parliament for the Dutch Liberal Democratic Party. She also curates the number one policy newsletter in the social media space that I read every week called Tech Policy Watch.

If you want to go deeper into the themes that we've been exploring in this episode and all the themes that we've been exploring on this podcast about how do we create more humane technology, I'd like to invite you to check out our free course, Foundations of Humane Technology at humanetech.com/course.

Your Undivided Attention is produced by the Center for Humane Technology, a nonprofit organization working to catalyze a humane future. Our senior producer is Julia Scott. Kirsten McMurray and Sarah McCrea are our associate producers. Mia Lobel is our consulting producer. Mixing on this episode by Jeff Sudekin. Original music and sound design by Ryan and Hays Holladay. And a special thanks to the whole Center for Humane Technology team for making this podcast possible. A very special thanks to our generous lead supporters, including the Omidyar Network, Craig Newmark Philanthropies, and the Evolve Foundation, among many others. You can find show notes, transcripts, and much more at humanetech.com. And if you made it all the way here, let me give one more thank you to you for giving us your undivided attention.