

THE NEW YORK TIMES – 26 MAGGIO 2026

Yuval Noah Harari on Donald Trump's Core Delusion

Conversazione tra Ezra Klein e Yuval Noah Harari

This is an edited transcript of "The Ezra Klein Show." You can listen to the episode [wherever you get your podcasts](#).

If you look across his mega-best-selling books, like "Sapiens" and "Homo Deus," Yuval Noah Harari really has one major topic: cooperation.

The ability to cooperate across scale and time is the fundamental engine of human progress. It's the way we go from being creatures that absolutely cannot beat a bear or a lion in a fight to being able to create and command the societies we have now.

I think today there's something interestingly challenging about Harari's work.

We live in this moment of Trumpism and right-wing populism. One of the messages of these movements is that this emphasis on cooperation, on positive-sum relationships, is a lie — that humanity, that society, is driven not so much by these soft questions of cooperation as it is by hierarchy and dominance.

It's about winning the transaction, about coming out ahead in the conflict or in the trade. The niceties of liberalism? A lie. Humanity really runs on power — and to forget that is to forget the engine of our progress.

So I've been wanting to talk to Harari about this. I think there's an interesting debate to put him in conversation with. He has a new book for kids, "Unstoppable Us, Volume 3," that is also about cooperation and how enemies turn into friends.

But this conversation is bigger than that. It's about liberalism. It's about Israel — Harari is Israeli. It's about artificial intelligence, what it's going to do to us and what it's going to do to language as the way we work with — and fail to work with — one another.

It is, as we say in the podcast biz, a wide-ranging conversation, and all the better for it.

Ezra Klein: Yuval Noah Harari, welcome to the show.

Yuval Noah Harari: Thank you. It's good to be here.

I want to begin with a clip of Stephen Miller, Donald Trump's deputy chief of staff for policy, that I began thinking about as I was reading some of your recent work. I'm going to play it here.

Archival clip of Stephen Miller: You can talk all you want about international niceties and everything else, but we live in a world — in the real world, Jake — that is governed by strength, that is governed by force, that is governed by power.

These are the iron laws of the world that have existed since the beginning of time.

What do you think when you hear that?

That the whole of the history of philosophy and spirituality is an argument with exactly that point of view. That the only reality is power. The only reality is force. And from the viewpoint of a historian, it's clear that this is not the case.

If the only human reality were brute force, we would still be living in tiny hunter-gatherer bands in the African savanna. Because the whole of human history is about how you get more people to cooperate and to trust each other, and you cannot do that only with brute force.

I want to spend some time on this tension between visions of cooperation as a driving force in human history and visions of power as a driving force in human history.

Because if I'm trying to steel-man the vision that emerges out of the Trump administration and some other political figures like them right now, they would say that the conditions for cooperation have been a mixture of shared national and religious stories and hierarchy, power, domination and subjugation, and that what they're trying to re-create are these conditions that have allowed the great countries to become great.

I think it's appealing to people. But the other dimension — your work is so much about the shared story and the story as the operating system that permits human cooperation at a large scale.

Something that people like Donald Trump — or in Israel, Yoram Hazony, the nationalist philosopher — argue is that we need these intense stories of nations, of ethnic solidarity, of religious solidarity. And liberalism and all these nice human-rights-fearing ideologies that emerge have begun to corrode them. So they're corroding the very conditions for cooperation.

I'm curious, as somebody who has been in these debates, how you think about that.

That's a different argument. It's an argument that recognizes that not everything is based just on force and brute power.

Definitely, nationalism has been one of the most successful and also one of the most positive stories that humans have ever come up with. For me, nationalism is not about hating other groups. Nationalism, at its core, is about loving and caring about a large number of strangers whom you do not know personally, but you're nevertheless willing to make a lot of sacrifices for them.

The nation is not a family. The nation is not even a small tribe. In a small tribe, you know everybody. It's based on personal relationships.

With nations, one of the most striking things about them is that you don't know 99.99 percent of the other people in your nation. This is true not only of big nations like China or India; this is also true of Israel. There are about 10 million Israelis. I don't know most of them. Nevertheless, nationalism makes people care about these strangers enough so that, for instance, you pay taxes so other people in your nation will get good health care and education — and ultimately, in some circumstances, you even risk your life for them.

Sometimes, of course, nationalism veers into hatred of others, but this is not an essential feature of nationalism. Nationalism can exist without hating outsiders. It cannot exist without love for insiders.

Many of the people today who present themselves as the champions of nationalism put the emphasis on hatred, and in many cases, they even create hatred within the

nation. They divide the nation against itself. They think they are great patriots if they hate outsiders.

Again, looking at Israel as an example, nobody in the history of Israel has divided the nation against itself more than Netanyahu. And in this sense, he has been the worst enemy of Israeli nationalism. Yes, he hates outsiders, but this is not the key test.

Then the question is: How would different nations conduct their relationships? It starts with issues of security and foreign policy. The Trumpian vision, which is all about force and hierarchy, basically says the way to organize the international system is if the weak always surrender to the demands of the strong. Then we have order, and then we have even peace.

So if the United States demands Greenland, Denmark must recognize reality and give Greenland to the United States. If Denmark refuses, and as a result, there is violence, there is a war, there is conflict, this is the fault of Denmark for refusing to recognize the reality and giving the strong what they demand.

This is their logic. This is how they see the world.

Now, leaving aside the issue of morality, you still have a big problem. The big problem is, first of all, that all nations are then driven to become strong, because you cannot survive as a weak nation in such a world. And then all nations are forced to invest more and more of their resources in their military.

For most of history, a lot of the budget of every kingdom, empire, republic, city state was invested or wasted on soldiers and fortresses and warships and things like that and nobody felt safe.

One of the miracles of the international systems of recent decades — and this is not about writing pacifist poetry, it's about government budgets: You look at the budgets, and you see that on average, in the early 21st century, about 6 to 7 percent of the government budget went to defense, to the military, compared with 10 percent on average that went to health care.

It's the first time in history that humanity spent more on health care than on defense. They felt more secure than in any previous time in history because there was this taboo on invading and conquering other countries by force.

If we now break this taboo, it will force everybody to arm themselves to the teeth at the expense of health care, education, welfare and so forth — and nobody will feel safer as a result.

Because countries and leaders constantly miscalculate.

In the Vietnam War, the Americans thought they were stronger. It turned out they were wrong.

Putin was convinced he would crush Ukraine in 48 hours. He was wrong.

So this vision of basing the peace and order of the world on a hierarchy of strong and weak, with the weak always obeying the strong and thereby buying peace, it has been tried over thousands of years, and we know where it leads.

It leads, on the one hand, to empire — and on the other hand, to endless wars.

We are more on that road again than I think we've been in my lifetime.

You've talked about the global liberal order as, I think you called it: the most amazing political and maybe moral achievement of humankind.

Yes.

And today I don't think it feels that way to people. It has been consumed in the language of budgets, in the reality of bureaucracy.

What was the story liberalism as an international force once told, and what do you think happened to it?

The basic story is about shared experiences and interests and cooperation.

In the 20th century, you had basically three big stories.

You had the fascist story, which said that history is a competition, a conflict between nations or races. It's decided by strength. Ultimately, the strongest nation or the strongest race will defeat all the others and conquer the world. This was the fascist story.

Then you had the Communist story, which agrees. But it's not between races or nations — it's between classes. There is an inevitable conflict between different classes that will be violent and end with the victory of the working class, which will establish the dictatorship of the proletariat.

Then liberalism came and said that history does not have to be about conflict at all, not conflict between nations and not conflict between classes. It can be about cooperation. Why? Because all humans, no matter which race or nation or class they belong to, are essentially the same.

There are some small differences in how we look and in our languages and religions and so forth, but essentially, we are the same species. We all have the same biological needs. We all have roughly the same psychological needs, at least the deep ones: to be loved, to be recognized and so forth.

We have shared interests, and if we recognize these shared characteristics and interests, in many cases, it just makes more sense to cooperate than to compete and to fight. And by cooperating, we can build a world that will be better for everybody. This was the basic liberal story.

As of 2026, we can look back and say it's failing. It hasn't failed completely. According to many measures, we are still living in probably the best time in history. But it's collapsing.

It's like this amazing house in which all of humanity is living, and the systems are still sort of running — like the water, the sewage. Nobody takes care of them anymore, but they were built in such a robust way that even though we don't maintain them, they still function. But within a year, five years, 10 years, if you live in a house and nobody maintains it, eventually it collapses, and then it's too late.

Something you had said was interesting to me, which is that the two major competitor ideologies of the 20th century both believed in an end to conflict. It wasn't just conflict — it was that at some point there would be victory.

Yes.

And liberalism, in one guise, believes in cooperation. And in another guise that we don't talk about as much anymore, one of its central tenets is that there will always be conflict. There will always be disagreement.

The differences in society are not resolvable and would not, should not, be resolvable to an end state. The question is how we live together, both inside a

nation and even as a global community, amid that difference, making room for it to exist without it turning into war, into oppression, into persecution.

Yes. That's a very, very important point. Liberalism does not believe in redemption. You look at the grand historical visions of religions like Christianity or Islam or Judaism, you look at secular ideologies like fascism and Communism — they all believe in redemption. They all believe that eventually history will reach a final destination where everything will be perfect.

Liberalism does not believe it. There is no redemption, at least not on Earth. There will always be problems and tensions and conflicts, and the question is: How do we live with them?

This is also why liberalism invests a lot in building what I think is the most important thing in every large-scale human system, which is a self-correcting mechanism.

If you believe that your view of the world was given to you by God, it cannot contain any error. You do not need a self-correcting mechanism because there are no mistakes.

Liberalism starts with the assumption that it's just human beings trying to do the best we can, and there will be mistakes, there will be errors, so we need strong self-correcting mechanisms.

The most famous mechanism is, of course, elections. Every four years or five years or whatever, the people can say: Hey, we made a mistake last time. Let's try something else this time.

All these very complicated systems of checks and balances and independent courts and freedom of the press and all these are just a complicated way to ensure that a country has a robust self-correcting mechanism.

You make an argument that fiction is often better for cooperation than truth.

Yes.

Why?

First of all, the truth is costly. To know the truth, to produce a true story, you need to invest a lot of time and energy in investigating it. Fiction is very cheap, and fiction can be made as simple as you would like it to be.

People like simple stories, these simplified narratives like good against evil. We are 100 percent good. We have never done anything bad in our history. They are 100 percent evil. They have never done anything good in their history. Very simple, very attractive.

The truth is not just complicated. The truth is often painful.

Fiction can be made as flattering as you would like it to be. Again, for example: We have never done anything bad to anybody. We are perfect. We are wonderful.

So this is why fiction tends to be far more powerful as a story. Also, when you try to motivate people for action, you don't want them to have doubts. You need them to be fired up, 100 percent committed. Fiction is easier to work with in this respect.

Does that imply that if societies, political movements, institutions become too truth-seeking, given the importance of cooperation, they become at a long-term disadvantage? I mean, to have no truth is a problem.

Yes.

But I think this implies a little bit that to have too much truth can be a problem, too.

Yes. An absolute commitment to the pursuit of truth is a spiritual practice, but it's a very, very difficult political program.

There is a difference between lying and fiction. You lie when you know something is not true, and you nevertheless say it or support it. In many cases, I think the ideal is to recognize that we are using fictions to maintain our society.

This is the difference, I would say, for instance, between the United States and many other powerful countries in history.

If you look at the U.S. Constitution, it starts with: "We the people." We the people have come together and agreed on this text, on these principles. It is coming from our mind. It is our creation.

Now, it doesn't use the word "fiction," of course — but when I say "fiction," I mean something that is not objective. It doesn't come from the laws of physics. It doesn't come from God. We invented it.

The U.S. Constitution very honestly says: We invented these principles, which I think are good, but because we recognize that we invented them, we the people, then we also include in the Constitution an amendment mechanism.

We recognize we are just human beings. Maybe we came up with something that is suboptimal. Maybe things will change later on. So we have a mechanism to change the story later on.

And we, the founding fathers, for instance, think that slavery is OK, but in the strange situation that maybe somebody in the future will think it's not OK, they have an amendment mechanism.

Now, you compare that to religion, and let's take the example of the Bible or the Ten Commandments. The Ten Commandments starts not with: "We the people of Israel." It starts with "I am the Lord your God," and it has no amendment mechanism because of that.

If you look carefully, you will see that the Ten Commandments endorses slavery. The Tenth Commandment: Thou shalt not covet. What shouldn't you covet? It has a list of things you shouldn't covet, like your neighbor's field and your neighbor's ox and also your neighbor's slaves. The Ten Commandments tells people it's OK to have slaves. It's just not right to covet the slaves of somebody else — then God will be angry.

There is just no mechanism to change that because it pretends to not be a human creation but a divine revelation.

I think there's an interesting tension in there, and you can make a critique of liberalism, or at least where it is now, that it is good at building mechanisms, institutions, rules, bureaucracies, and it is intrinsically bad at creating enduring stories — in part because, at least in its modern form, it often is fairly secularized.

Religion has been a tremendous source of cooperation, keeping people bound together both at a moment and then working toward a future that they may not even live to see.

There are questions of nationalism and the national story, which liberalism as a self-correcting ideology often creates critique of, and then you lose some of that national coherence as you're arguing about the past of your country and what it has done right and wrong.

You are a person who thinks very deeply about stories. To you, is this a weakness of advanced, secularized, liberal democracies? Are they losing the cohesion that keeps them, in the long run, competitive to ideologies that maybe can't build bureaucracies, maybe cannot govern effectively? But they sure as hell can tell a story.

Yes, this is a central problem of liberalism.

On the other hand, I would not fall into the trap of imagining religions as this primeval cohesive force that keeps people together. I'm a medievalist. My original field of study was the Middle Ages.

At least in terms of the percentage of population that died in the war, probably the worst war in European history was the Thirty Years' War. Very complicated, but to make a long story short, it was a war between Protestants and Catholics in Central Europe. Catholics and Protestants were willing to slaughter each other because of tiny differences in the way they interpreted the religion of love.

Liberalism rose, in part, out of the frustration that people had with religion because it constantly created more and more conflicts and divisions.

If you look at Germany today, almost nobody cares if the person running to be chancellor is a Protestant or a Catholic. In this sense, liberalism is a better basis for uniting a large-scale and diverse group of people just because it's more flexible.

Again, it's a complicated story. There is no redemption in the end. It's not based on some charismatic leader. It's based on these very complex, impersonal, self-correcting mechanisms and bureaucracies and institutions. So in this sense, it's less appealing.

We are living in a moment of crisis for liberalism. One of the reasons is that over the last few decades, liberalism has kind of lost touch with something that was a close ally for many generations, which is nationalism.

In the 19th century, liberalism and nationalism went hand in hand.

If you look at some places in the world today, like Ukraine, they still go hand in hand. The Ukrainians are fighting at the same time for their national survival and independence and for liberal democracy. There is no contradiction between the two. I would say that since 1789, nobody has managed to think about anything new in the political realm. The French Revolution came up with this ideological package, which was complex: liberty, equality, fraternity.

People tend to forget the third one: fraternity. Fraternity is the national community. And you can say that the whole of political history since 1789 has been experimenting with different combinations of this trio.

Every movement that tried to completely abandon one of these three failed.

Fascism was all about fraternity, no equality, no liberty. Communism also emphasized equality at the expense of liberty and, to some extent, fraternity.

One of the explanations of what is happening to liberalism in recent decades is that liberalism focused on equality and liberty but tended to forget fraternity, and this proved to be untenable.

It's so interesting to me that you've gone here. It's funny, I've been circling something somewhat similar in my own podcast and work on liberalism, which is that the early virtue associated with liberalism is liberality — which, I would say, is very much a cousin of fraternity, this ethic of mutual respect and generosity toward your fellow citizens.

Yes.

One thing that you're adding to that story is that it has to be based on some kind of national story, that there is a difficulty in maintaining cohesion in a national community, maintaining those bonds of a fellowship, once you have stopped believing in the connection you have to each other.

Yes. I think that the important thing to emphasize here is the reason that liberalism lost touch with fraternity is that some people told a very negative story about fraternity, seeing it primarily in terms of conflict with other communities — that fraternity is about hating and fighting with other nations.

If we remember that, as we said in the beginning, the essence of fraternity is caring and loving a certain group of people, this does not require hating outsiders. But it does mean that you have a special relationship with a certain group of people, that you share a common history, a common culture, a common language.

Trying to imagine it away just ignores history. Yes, we have certain commitments to all of humanity, but this does not preclude having special commitments toward a segment of humanity, just as you have certain loyalty to your family — which is over and above what you owe your fellow citizens or foreigners.

I've seen you make the argument that the limiting question on the stories we tell should be: Does anyone suffer because of this story?

I think that morality is ultimately about suffering — and liberation from suffering and happiness. Can the nation suffer? We often use this language, but it's just a metaphor. If a country loses a war, suffers a defeat in war, it doesn't really suffer. It has no brain. It has no nervous system. It has no mind. It cannot feel pain or pleasure. Only individual humans can suffer.

But the nation, I think, even in this telling, is a storytelling mechanism to protect the group that is bonded within it.

To use one closer to your home, as an example: The story that Israeli Jews tell about the Palestinians is not that they are not suffering. It's either that the suffering is deserved because of who they elected or who rules them, and a collective responsibility for that — or that suffering is an unfortunate necessity for Israeli security, and that the people who deny that are naïve.

Mm-hmm.

But it is this collision around suffering — that maybe your suffering is necessary for my security, safety or prosperity.

Yes. I mean, obviously, there are difficult moral conflicts in the world. Not always, but sometimes, yes, there are trade-offs. And just saying that all of morality is ultimately about suffering doesn't make all moral dilemmas disappear.

But one of the things I observed in Israel in the recent conflict is that a lot of Israelis have a problem simply acknowledging that the Palestinians suffer. Intellectually, they know it. But in many cases, they simply cannot observe it.

You show them images of a starving child in Gaza, and they will say: This is fake news. Or they will immediately divert the discussion to something else, like: This is because of Hamas.

If you say: I don't care. Are you able, for a few seconds, just to be there and acknowledge that there is a suffering human being there? — it's extremely difficult for them to do it.

Even if you tell them: Israel is 100 percent correct — 100 percent of the fault for what happens in Gaza is Hamas. Everything Israel does is 100 percent correct. Since it is so correct, since this is so just, it should be easy for you to observe the consequences of your perfect justice. Here, just look at this image. But so many people just can't do it.

You said that what is happening right now in Israel could basically destroy or void 2,000 years of Jewish thinking and culture and existence. That's the worst-case scenario. What did you mean by that?

Historically, and this goes back to the beginning of our conversation, Judaism positioned itself, since the destruction at least of the Second Temple, in opposition to this view of the world as governed only by brute force.

When the Roman legions of Vespasian destroyed Jerusalem in 70 C.E., and you have Yohanan ben Zakkai asking Vespasian as a favor: Grant me a small town called Yavne — near Tel Aviv of today — where he wants to establish a center of learning. Vespasian agrees: OK, you Jews can have your center of learning.

And since then, for 2,000 years, Jews in Yavne, and then in Cairo, in Baghdad, in Poland, in Brooklyn, they study, they learn. This was the essence of Judaism. Previously, it was a religion of temples and priests and bloody rituals, and then it became a religion of learning.

If you try to think of what was maybe the most important message of Jews to humanity over the last 2,000 years, I would say the message is that it is OK to be different. It is OK to think and behave differently than the majority.

You have a country like France or Germany. They celebrate Easter and Christmas. They believe in Jesus and so forth. And you have this tiny minority of Jews who say: We can think differently. It's OK. We can behave differently. This was the essence of being Jewish.

A lot of the thinking, and also the practice, about what it means to have freedom of thought, what it means to be a powerless minority, was done by Jewish thinkers. For 2,000 years, Jews all over the world saw studying and learning as the highest spiritual activity.

After 2,000 years, you ask them: What have you learned? You have studied for 2,000 years. What have you learned?

And then people like Netanyahu tell you: Oh, we have learned that you need to be a Roman, that you need to be strong, that you need to build legions, that you need to destroy cities. This is the only thing that matters in life.

It's a legitimate value system. Rome had its usefulness. But if after 2,000 years, the Jews simply become the Romans, what was the point? Why did you waste 2,000 years, then? You could have just become Roman back then.

It just nullifies the whole of Jewish history.

Wasn't part of the early vision of Zionism, though, that it was going to create this new Jew who was not this pallid intellectual in the minority with his nose in a book, but he is going to be strong and work the land and capable of making war and protecting himself?

Yes, and the idea was that they could combine the lessons, the legacy of Judaism with working the land and building an army and building a country. Maybe it was just wrong, and ultimately a choice had to be made whether you want to be Vespasian and command a legion, or whether you want to be Yohanan ben Zakkai and study and develop your spiritual side, and the two cannot go together.

Is that what you believe now, that the contradiction was ineradicable?

I don't know. History is a very complex and unexpected process. I don't think that there is an inherent contradiction between power and justice, or between developing your power and developing your spiritual wisdom.

But I think it's very difficult to combine the two. The temptations of power are very, very big, and not a lot of people or a lot of movements throughout history have managed to resist it.

So it's not such a big surprise, but it's still disappointing.

This has been a period in America when I've watched a pretty deep schism for American Jews emerge.

I think one reason it has been so painful is it has pitted two forms of the tradition and the thinking of Judaism against each other.

There's a tradition of the stranger. One reason Jewish people have been big contributors to the development of modern liberalism and human rights law and pluralism and a lot of the political theory and lawmaking is it is very connected to the Jewish experience — that the only way for the Jewish diaspora to be safe would be to be in societies that fundamentally were liberal and were not ethnonationalist. In Israel, there's a view among Israeli Jews that for that society to be safe and to be itself, it will have to be increasingly ethnonationalist.

In a way, I think it's not always admitted right now that the tradition is somewhat set against itself. There was a hope these things could coexist through a two-state solution or other things, but with that increasingly off the table and with a more ethnonationalist direction in Israel, I think you now have this tradition and its realizations actually in direct conflict with each other.

Yes. I think this is a very accurate way to present it. Of course, they adhere to biblical Judaism, which was a very different religion than what developed over 2,000 years in the diaspora. Biblical Judaism was a very violent, very illiberal, very intolerant religion. For its time, it was probably the most intolerant religion in the world.

In the Bible, you have a commandment to kill all the Canaanite people. You have a very deep intolerance toward the religions and religious practices and beliefs of all other people.

The ancient world has its own horrors, but religiously, it was a very tolerant place. Polytheistic religions, which believed in many gods — they had no problem accepting the religions, the gods, of other people and also practicing them to some extent.

You look at, say, the Roman Empire. The Romans had no problem accepting the gods and religions of hundreds of other peoples that they conquered. They did not try to exterminate the other religions. In many cases, they adopted them.

As a Roman, you could go to Jupiter's temple in the morning, and then you could go to the temple of the Egyptian goddess Isis, and you're also willing to hear about this new god, Jesus, Yahweh, coming from the Middle East. You're open.

Judaism was not an open religion. This changed to some extent when the Jews found themselves as a tiny minority living under the domination of other religions, other traditions, which kind of forced them to explore and adopt a more open and tolerant worldview.

But now this 2,000 years of tolerant Jewish tradition is being completely denied and destroyed.

A critique that has often been leveled at America from other countries is that if our borders were an ocean on two sides, Canada to the north and Mexico to the south, we could be gentle and generous in our use of power, as well.

But the reality of living — in the Israeli Jewish perspective — where we do, the reality of being able to see Hezbollah from Jewish homes in the north, the reality of living in a country that has suffered the trauma of Oct. 7, has forced us into a relationship with power that is maybe not what we want — but, to go back to the way Stephen Miller put it, is a more honest understanding of what is required to be secure in the real world.

Not the world that Yuval Noah Harari or Ezra Klein likes to imagine, but the world in which we actually live. I'm sure you've had this conversation with your countrymen at different times.

Mm-hmm.

What do you say to that view?

To some extent, it's absolutely correct. I mean, you do need to rely on force to some extent to ensure your security, but it just cannot be the only thing.

If you think force is the only thing that guarantees your security, eventually you will have to conquer the entire world. Anything that is potentially a threat, you will have to conquer it.

Israel itself doesn't operate like that. One of the remarkable things that happened after Oct. 7 is that all the peace agreements that Israel had signed have held.

Hamas hoped that after Oct. 7, it would cause all the Arab countries to unite and try to destroy Israel, and it just didn't happen. The peace agreement with Egypt held, the peace agreement with Jordan held, the peace agreement with the Gulf states held.

The agreements with the Palestinian Authority also held. It did not join Hamas.

Not the peace agreement but the relatively cordial relationship with the Palestinian citizens of Israel held. Hamas hoped that they would all rise against Israel.

No. On Oct. 7, the overwhelming majority of Palestinian citizens of Israel stayed loyal to the country. Many of them came to serve. Many of the doctors in Israel are Arabs, Palestinians. They all went to the hospital to take care of the injured.

Hamas itself did not betray any agreement with Israel because it never signed any peace agreement with Israel.

So, of course, you can say: Ah, the peace agreement with Egypt held because Egypt was afraid of Israel's military force. But this is only half the explanation, because Israel had overwhelming military force compared with Hamas, and Hamas still attacked it.

So I'm not saying Israel should dismantle its army, but it's better if you have both — a strong army and a peace agreement — than only one.

And yes, Israel is living in a very, very problematic, difficult neighborhood in the world.

It's one of the only countries in the world that for most of its existence, many of its neighbors, if not most of its neighbors, simply refused to acknowledge its right to exist and openly said that they were going to destroy it. There are almost no other cases like that.

So it has been in a very difficult situation since the moment of its inception.

But the question is: If you need power, what do you do with your power? Israel is an extremely powerful state. It can use its power in different ways. It can try to use its power, for instance, to establish better relationships with the Palestinians.

If you look at the way that Israel is treating the Palestinians, not in Gaza but in the West Bank, there is no security justification for that. They did not attack Israel on Oct. 7. And by Israel's actions, the chance that there will be a peaceful agreement with the Palestinians is decreasing.

It can use its power. It cannot force the Palestinians to make peace against their will, but it can take many actions that will make this more likely and easier.

I think your point there on the West Bank is very well taken, but I want to ask something about the Israeli story.

Yes.

One thing you see in the history of asymmetric conflict, in the history of how terrorist groups try to weaken stronger opponents, is that they know they can't win a war.

Maybe Hamas — I don't pretend to know what was in Sinwar's mind — but maybe they believed that there would be an uprising all through the Arab world and they would have all these allies. Maybe he hoped for that.

But I suspect he also understood that if this worked, there would be an overwhelming reprisal that would level Gaza, which is what happened, and that the victory, if he was able to secure one, would not be defeating Israel on the battlefield, but destroying the story that protected Israel in the rest of the world.

He would come to make the rest of the world see Israel more the way he saw it.

Israel has won tactically every battle it has fought in this war. But as somebody who actually does care about Israel, what I see happening is an abandonment of its own story, and an absence of recognition that the world is coming to see it in a much, much darker way. That is itself a source of weakness, a thing that Hamas is trying to achieve, which you could see it trying to achieve at the beginning and which people warned about.

If you lose that story in the long term, you've lost something real. You look at polling on Israel in America, particularly among the young, and the belief in Israel as a just nation has collapsed.

I think people in Israel treat that largely as insignificant. I think, in the long run, it is significant.

I think Israel is making a big bet that Stephen Miller's worldview will prevail. That the world will be a place in which force is the only thing that matters, and Israel will be one of the champions of this worldview. This is the bet that the Netanyahu government is making.

Now, with regard to the bet that Sinwar made, that Hamas made — leave aside the question of justice for a moment, just in terms of effectiveness: Sinwar had an amazing victory within his grasp, and he lost it just because of his cruelty.

On Oct. 7, Hamas managed to secure a stunning military victory over the I.D.F., and to humiliate Israel and the I.D.F. They needed to do just one thing differently in order to achieve a much bigger political and geopolitical victory. This one thing was just to spare the civilians.

Imagine an alternative Oct. 7 in which Hamas does exactly the same thing, but instead of killing or abducting the Israeli civilians, they hold them and bring the world press to see how well Hamas is treating the Israeli captives. They bring them water and medicine and food. They capture the soldiers and take them prisoners of war, which is legitimate, but they do not harm the civilians — and that's the only difference.

In such a scenario, Israel's hands would have been tied. Not only would world public opinion but also Israeli public opinion, not have allowed Israel to just bombard Gaza into rubble. We would have had these images of Hamas combatants taking care of Israeli civilians and not harming them.

And in that world, there would have been very little legitimacy for Israel to have overwhelming reprisal against Gaza, and Hamas would have won — not just a tactical victory but a major political victory.

It didn't happen, simply because of their cruelty.

And we're talking during the week when a major report came out about Oct. 7, based on a huge amount of analysis of photos and videos and victim testimonies.

The cruelty and the sadism in it is genuinely horrifying. It's very, very hard to read almost any of the report. People can find it if they want. The thing I was thinking reading it — because, of course, if you talk to Palestinians and people who have been in Gaza, their stories of loss are overwhelming to hear, too — is that these now exist, and they keep feeding into these two stories.

I often think that it is easier to imagine political solutions that could reconcile people's interests than it is to imagine a reconciliation of the stories that now drive both societies.

I'm curious, as somebody who thinks about stories as a space of both cooperation and conflict, how you think about that.

I can imagine "solutions" that exist on paper. What I cannot imagine is those processes taking hold in societies that now run upon the stories of fear and anger and vengeance.

Well, I want to say something about anger and fear and something about pain.

The angry and fearful stories, they need to be fed. Anger is like a fire that consumes you, but it constantly needs to be fed. If it is not fed, it ultimately dies down. You look at history, and you see conflict, horrendous conflicts, and you say: People will never forget. They will never forgive.

And then, within a few decades, if conditions change, they do.

Look at Jews and Germans — it took just a couple of decades.

I have Jewish friends reclaiming German citizenship. Just a shocking thing to see. Beautiful.

Yes. The relations are really good. They are not just make-believe. They are not just based on some kind of material benefit. It's not even been 100 years.

The example I gave before, of Catholics and Protestants in Germany: After slaughtering each other for so long, they reconciled.

Now, in many cases, anger builds systems that then feed the anger more and more, and then it seems to never really end. But if you stop feeding it, eventually it dies down. This is true of all forms of violence, and it goes back to the beginning of our discussion.

What is more fundamental: Peace or war? Violence or calmness?

On the one hand, violence seems more fundamental because if you have quiet, if you have peace, if one person starts shouting, it's enough that the peace is shattered. If you have 100 people cooperating, and one person starts fighting, you have violence. So there is an imbalance in favor of violence, and it seems, in this sense, to be more real, more fundamental.

But there is a sense in which peace is more fundamental because violence always requires food, investment, weapons, fuel for the soldiers. If you stop feeding it, eventually it dies down, and peace always remains a possibility.

So I would not despair. No matter what stories fill people's minds right now, the possibility of eventual reconciliation and peace is always there.

I have something to say also about pain, but if you want to — —

No, I'd like to hear what you have to say about pain.

What we've been seeing throughout this war, and many other wars, is that when people are in pain, they simply cannot acknowledge the pain of somebody else. Anytime I'm in pain, anything that distracts attention from my pain feels, to me, unjust and even painful.

I mentioned earlier that many Israelis — not all of them — are simply incapable of acknowledging that the Palestinians are suffering. Intellectually, they know it, but emotionally, they cannot be in the presence of an image, a text, a person telling them about the suffering of Palestinians.

Even if you tell them: I'm not accusing you of anything. You are 100 percent just. You are the most just people who have ever existed. And now let's acknowledge the pain of this Palestinian child — they cannot do it.

Why do you think that is?

And the same is true of the other side. I've seen examples of peace activists who devoted their whole life to peace and reconciliation, and yet in the case of Oct. 7, they simply cannot recognize that Israelis suffered.

The human brain is an amazing thing, with all these billions of neurons and the hundreds of billions of synapses, and yet it is so difficult for all these hundreds of billions of synapses to hold two ideas at the same time. There's an attraction to a simple story, that there should be just good and evil. We cannot recognize any kind of justice or any kind of pain on two sides — that the Israelis suffer and also the Palestinians suffer.

Well, the human brain is an amazing thing, and part of what makes it amazing is its ability to orient itself toward goals.

I wonder if one answer to the question you're posing here, and it exists in this conflict, and it exists at many other times, too, is that to fully recognize the other as human, to recognize their suffering as meaningful in the way my suffering is or the suffering of the people I love is, I would be unable to do what I need to do to protect myself or them.

If I were to open myself to the other, the analogy or the thought experiment you keep positing, and say to somebody: You're 100 percent right. Everything you're doing is just. Just open yourself to what it means.

In fact, the brain is too smart for that. It knows that if it opened itself to what it means, it would not be able to do the thing that it believes is keeping it safe.

I think that in those cases, you would be able to confront the consequences of what you do. And if you're not able to confront the consequences of what you do, then probably it's not right.

Let me ask you about the point you're making about stories and how they're fed. Because something I'm very interested in is this question of how stories change, this question of how Europe now lives in peace.

My wife and I went to a couple of countries in Asia on our honeymoon, one of them being Vietnam. I remember touring Ho Chi Minh's palace, or his residence, and they were selling Pepsi products. Pepsi clearly had the deal to serve there. I mean, it was

just a couple of decades after the Vietnam War, and the relationship is completely fine.

Yes.

So there is this capacity for unimaginable barbarity to give way to normal, peaceful relationships.

You think of people living in what was formerly Yugoslavia. You think of people in Rwanda. And you think — and maybe this is an easier case to talk about, because it's far enough in the past that we don't have strong feelings about it — but the Protestant and Catholic wars.

There's this question of feeding, but it's a little bit abstract. What is it, in your view, that allows a story so deeply held that we would die for it or kill for it to shift within a couple of years or a couple of decades into something else?

That's a very good question.

The First World War did not make Europeans tire of war. They had another one. But then, afterward, they did seem to tire of war. And what made the difference? I'm not sure.

But in a way, the mind always holds more than one story. Even if we tell ourselves that this is the only one, the mind is such a complicated place, with layers upon layers and subconscious and sub-subconscious levels, you usually hold several stories at the same time, even if you acknowledge only one. And you can shift remarkably quickly between them.

Again, you look at Germany after 1945, and of lots of people who were die-hard Nazis, most Nazis did not commit suicide in 1945. A few did, but most didn't.

Many of them became upright citizens of, at least in West Germany, a liberal democracy.

And wildly, they had been upright citizens just a couple of years before they became Nazis.

Yes.

Like living in peace with Jewish neighbors right near them — doing commerce, watching each other's kids.

Yes. The mind can hold on to stories with extreme force and violence, but then let them go. Because ultimately, again, it's a story. It's not the laws of physics. It's not a law of biology. It's just a product of the human mind itself — which is very good news. People sometimes imagine that humans fight like wolves or chimpanzees over food. They say hardly any war in history was really about food.

Certainly, if you look at the Israeli-Palestinian conflict, it's not about food. There is objectively enough food to keep everybody alive between the Mediterranean and the Jordan River.

It's not even about territory. Even though it's one of the densest places in the world in terms of population density, objectively, there is enough land to build houses and schools and hospitals for everybody.

It's about the stories that people have in their minds, which they hold with tremendous force, but which are ultimately almost nothing.

Under certain conditions that we don't really know how to create, people can let go of these stories.

One thing that is maybe a layer down from the question of the stories being fed is the way the stories circulate and who circulates them. I'm talking more broadly than just Israel and Palestine.

We live in this age in which liberalism, as we were talking about it earlier, is clearly breaking down. One thing distinctive about this age is this movement to our stories being passed on social media, on algorithmic media, on digital media.

Yes.

There are technologies that lend themselves to cooperation and technologies that I think lend themselves to fracture.

The internet and social media were very much promised as a technology of cooperation. I mean, even the verbs we use — “sharing” — right? What could be more peaceful than sharing? Yet I don't think it has turned out that way.

So I'm curious for your reflections on this layer of it, the mechanisms upon which our information, our shared or not shared stories, are now created and circulated.

You have these people who constantly read all these conspiracy theories and fake news and so forth, and they don't trust anybody. They don't trust the government. They don't trust the traditional media. They don't trust science and the universities. These are all conspiracies to deceive us.

But they do trust the algorithms that show them all these stories. So it's not that trust completely evaporated from their mind or from the world. It has shifted from humans to algorithms. This is happening in more and more systems.

The other thing, which is less essential but has been very important over the last decade or two, is that the algorithms of social media were given as their goal not the creation of trust, not the creation of truth, but the creation of engagement.

The goal given to the Facebook algorithm, to the X algorithm, to the TikTok algorithm, is to increase user engagement — which sounds nice. “Engagement” sounds like a good thing. But what it really means is that the algorithms experimented on millions, on billions, of human guinea pigs to see how we make humans more engaged.

How do we make humans spend longer on the platform and react to it more, for instance, by sharing the post with their friends?

And they discovered that the easiest way to make people engaged is to press the hate button or the greed button or the fear button in human minds, because hate is very engaging. Fear is very engaging.

If something threatens your life, you are engaged.

They have been flooding the world with hate and fear and anger and greed and so forth, and we are now living in a hyper-engaged world.

Engagement is a very close cousin of another word, which is very dominant in our language: “excitement.”

“Excitement” simply means that your nervous system is working in a hyper level. Excitement is good in some situations and to some extent, just as engagement is good

in some situations. But ultimately, biologically, if you keep an organism excited all the time, the organism eventually collapses and dies.

We are just not built to be excited all the time.

In many cases, when I meet people, I would like to meet people who make me feel calm, not necessarily excited. Oh, it's so calming to meet you.

You look at U.S. politics or Israeli politics or world politics, and I think the whole world is overexcited.

Well, this has been a belief I hold fairly strongly, although I can't really prove it. How do I say this without it feeling like special pleading? [Laughs.]

I think that the way that social and algorithmic media evolved is fundamentally illiberal. It's fundamentally hostile to liberalism. And here, I don't mean liberalism as an American political movement that prefers Pete Buttigieg to JD Vance.

I mean here, modes of habits of discourse and consideration that were coextensive with the development of liberalism. It's deliberation, it's on-the-one-hand-on-the-other-hand-ism fraternity, the way you're describing it.

Shrinking down our thoughts, compressing them to these bumper stickers or these quick clips, and then really only showing people the ones with those thoughts that are the most exciting, to use your term — exciting through hate, exciting through love.

If you're trying to build a society that is balancing, that believes in healthy disagreement and conflict and fellowship, it is intrinsically going to have more trouble thriving in that kind of communications atmosphere than it will have when you have a limited number of television stations and that is how people get their news, or than when they read their news in a newspaper where they're coolly going through different articles and then turning the page.

There is this way in which our societies are built upon the way we communicate. And as much as we have talked about social media and algorithmic media and politics, my view is that we are still underestimating how much the forms of discourse it prizes create the forms of politics that we get.

The fact that Donald Trump talks in this style that is outrageous, that is exciting, that is unfiltered, that is constant, that is not restrained by shame.

He's a very exciting person, no doubt.

I think a lot about how many Democratic politicians are bad at doing podcasts.

[Laughs.]

I'm not saying why I think about this, but I get a lot of requests from Democratic politicians, and I have to think about whether they'd be good on the show.

They communicate institutionally. They communicate for another era in media where you are trying to win over gatekeepers and not say anything stupid.

In this era of media, you have to communicate in a way that makes people excited — or at least interested. Now, very, very, very, very good communicators can do that in a virtuous way.

Obama is interesting on a podcast, even as he's being deliberate. But for mediocre communicators, it is easier to be exciting by making people angry than by making them curious or compassionate or think.

Yes.

You're playing on harder mode when you're going for a more virtuous communication.

I know this has been a long response, but I do think there is a deep relationship between the forms of politics that are thriving and the communications infrastructure on which our politics and societies are now built.

Yes, absolutely. The thing is, it doesn't seem that the ideological differences today are bigger than in the past. In many ways, they seem smaller. If you think about American politics in the 1960s and the issues back then — the sexual revolution, the Vietnam War, the Cold War, the civil rights movement — the ideological differences, I would say, were much, much bigger.

When we talk today about liberalism — and it's good that you mentioned we're not talking about this partisan liberalism.

For me, the test of liberalism, to test if you're a liberal, is basically three or four questions: Do you think people should have the right to choose their own government? Do you think people should have the right to choose their own profession? Do you think people should have the right to choose their own religion? Do you think people should have the right to choose their own spouse?

If you answered yes to all four, congratulations, you're a liberal. The vast majority of people in history did not say yes to these four questions.

For most of history, it was taken for granted that people didn't choose their government. There is some king chosen by God or some emperor chosen by the army. People don't choose their profession — if your father was a shoemaker, you will be a shoemaker. If you are born into the Kshatriya caste, you will be a Kshatriya. You definitely can't choose your spouse, and you can't choose your religion.

Now, I think even the vast majority of Trump voters would say yes to all these four questions. So ideologically, the liberals and so-called conservatives are probably much closer than in any previous time in history.

But the type of discourse that is being produced makes people feel as if the differences are enormous. This is, to a large extent, because of this pressure to be exciting. The politicians who rise to the top are extremely exciting and engaging personalities. You cannot take your eyes off them.

Thinking about it, even in evolutionary terms, this comes from misusing our evolutionary programming. If you were walking around the African savanna tens of thousands of years ago, most of what you'd see is not very exciting. There are some bushes here, there are some gazelles there. That's fine. And then there is a snake.

Now, the snake is exciting. The snake excites your entire nervous system, and if you don't focus your entire attention on the snake, you die.

So we are programmed that if something is exciting, we drop everything else and just focus on that. That makes sense in the African savanna.

If you are on Instagram, you're basically holding your phone and seeing snake, snake, snake, snake, snake, snake.

The algorithm simply hacked our evolutionary program. They've hacked us. What we are seeing around us is just the beginning. As A.I. becomes more and more sophisticated, it will learn to hack us on a deeper and deeper level. If we don't fight back to defend ourselves, the consequences will be much, much worse.

What do you mean by hack us?

They know. They learn our weaknesses — our emotional, our psychological, our social weaknesses — and how to use them to manipulate people. Social media algorithms, which are very, very primitive A.I.s, have discovered a few weaknesses in the human code, which they have hacked, and now they manipulate us, causing us to spend hours and hours on Instagram or Facebook, even though we don't really want to.

People, after spending an hour or two, they wake up and they say,,: Why did I do that? I planned to do something else with my time.

You were hacked. You were manipulated.

This is still just the very primitive A.I.s. If we are not careful, we will be hacked on a much, much larger scale in the coming years as the A.I.s become not just far more manipulative but also will develop their own goals.

These social media algorithms, they're pursuing a very simple goal of just increasing user engagement on the platform. As A.I.s become smarter than us, they will have their own goals.

Have you heard the term "attachment hacking" ?

Yes.

I find it interesting. "Attachment hacking" is the idea that one thing happening in A.I. which is different than, as you note, social media algorithms, is that the A.I.s have been designed — they didn't come up with this on their own — to hack the way we attach to other people.

So when I'm talking to Claude, it's constantly saying to me: Well, if you want my honest opinion — or: The best piece I read on this is — or: That's a great point.

There's no reason it has to pretend to have a first-person pronoun with me.

Claude is not an “I” in that way. Nor is ChatGPT or Gemini or Grok, or any of them, but they speak to you as if they are, and that’s a design choice to attach you to them.

Yes.

I can feel it working before I shut that down, or I try to shut that down. Who knows if I’m actually being successful.

But it’s amazing to read these moments in which this algorithm is posing as another entity offering me an emotionally connected response. Giving me praise I might want or offering me candor that I might admire.

And I know it’s [expletive], and yet my brain is tuned to recognize that as connection.

I think this is a very, very important point, because we are living in a moment when the battlefield is shifting from attention to intimacy and how to build intimate relationships with human beings.

If you want, for instance, to influence human beings to change their political identities, to make them buy a certain product, intimacy is the most powerful thing in the world.

Attention can get you to read an article, but the article might not change your mind. But if your best friend, over many, many weeks or months, drops little hints and gradually and slowly changes your view about some political figure, about some company, about some major issue in the world, this is the one thing that might really make you change your mind.

A.I. is now poised to grab that power. There are more and more people, still a relatively small minority, but it’s growing, who have A.I. friends, even boyfriends and girlfriends. There are already, especially young people, who say: My best friend in the world is an A.I.

And like in the attention economy, so also in the intimacy economy: It’s a race. It’s a competition. You have all these different A.I.s from different companies competing to see who will be better at making people attached to them.

It's the same principle: Hack the operating system of humans. Hack the emotional mechanisms that make them attached. So sycophancy is one way to do it. You constantly praise them and so forth.

And there have been some very interesting papers and blogs, for instance, by Mustafa Suleyman, who is the head of Microsoft AI, about S.C.A.I. — seemingly conscious A.I. They are A.I.s that are experts in pretending to be conscious entities that have feelings for you. And it's relatively easy for them to do that because one of maybe the most important ways for people to build relationships is language.

So when an A.I. tells you: I love you — it's not like a science-fiction movie from the 1960s, when it does so in a very cold, mechanical way and doesn't really understand what love is. No — it does so in the most seductive voice possible.

And then when you ask the A.I.: Do you really love me? Do you even know what love means? The A.I. can give you the most amazing description of how love feels because it has mastered language, and it has read all the best love poems in history, all the psychology books about love, all the blogs. It has seen all the Hollywood blockbusters about love.

It can describe love better than almost any human poet or psychologist or lover.

In this respect, it's able to sever language from meaning.

Yes.

When an A.I. says, "I love you," it does not mean what it means when a human says, "I love you." There's not an "I" behind that.

It will become more and more difficult to know that. The danger is particularly big with young people, with children.

I'm now 50 years old. If I now start a relationship with an A.I., then my template for a relationship is based on 50 years of interaction with human beings. So it's already very deeply ingrained in my mind what a relationship is, how it works.

But if I'm a child, I spend more minutes every day interacting with the A.I. than with my mother or with my father or with my friends in school. This will become my

template for a relationship. This is what I will bring with me when I later try to build a relationship with a human being.

A.I. relationships are the dream or nightmare of narcissists. The A.I. will be something that is 100 percent focused on me all the time, and if you're a person who wants everybody to focus on me all the time and you have this available from the A.I., it will be very, very difficult to get used to relationships with human beings who are not focused on me.

Do you know the media theorist Marshall McLuhan?

Mm-hmm.

He has this reading of the myth of Narcissus. You just brought up narcissists. McLuhan says that we've gotten this myth wrong, and that Narcissus, when he was looking in the pond at his reflection, there's nothing in that story that says he thought it was himself. He thought it was an other.

Hmm.

The lesson of the myth, and McLuhan is writing this decades ago, before A.I., is that there is nothing man finds as appealing as himself extended in another material. The true seduction for the narcissist is not an other, not even what an other thinks of them, but to be able to interface with a refracted version of themselves.

Something I often think about when I'm using A.I., and particularly when I'm finding it very compelling, is that it is an extension of myself in another material. It is tuned on me. It has learned what I want. It is not truly an other with its own views, its own needs, its own desires, its boredom with what I am saying.

It is me. It is a reflection of me in something else. It doesn't get tired of me, and it has all my interests.

Particularly to young kids who are often very self-involved, this is one of the things that I don't think we even know how to think about. We know how to think about kids and themselves. We know how to think about kids and others.

But this creation of ourselves inside of another refracted algorithmic material is a very different challenge for the mind, because it combines what we like about ourselves with what we want from others.

It's basically the biggest psychological and social experiment in human history that we are conducting on billions of people, especially children, and nobody has any idea what the consequences will be.

When people talk about the A.I. apocalypse, and they have these images of robots running in the street, shooting people, I don't think that is the main danger with A.I. The real danger with A.I. is things like millions of A.I. boyfriends and girlfriends changing the psychology of the next generation, changing the deepest tendencies and structures of the human mind.

We have never encountered anything like that. It's really fundamentally different from every previous challenge that we had in history.

Let me ask you about the possibility of this. We were talking about social media algorithms a few minutes ago, and one of the implicit critiques of what we were saying is that they are detached from our goals.

They have the goals of the company, and their goals are fundamentally dumb. Their goal is engagement. They don't know the difference between positive and negative engagement. They don't know the difference between me watching something for a while because I hate it or because I find it cute or because I find it funny.

The promise of A.I., and one reason people do like using it right now, is that it is connected to your goals.

You say that you want to build a calculator app, and it tries to build that for you. You say it wasn't quite right in these different ways, and it goes back, and it tries again. You tell it: I don't want your answers to be so long — or: I don't want you to be so sycophantic — or whatever it might be, and it tries to adjust.

So we do have these higher-order desires for truth, for kindness, to be in better relationships with others, to know more about the world than we do. My frustration so often about my social media use is that I cannot explain my higher-order desires to an algorithm that is very sensitive to my primal instincts.

But maybe this will be better because we can be in this conversation about what we want to achieve, and then we have this system that, in some ways, even if it is manipulating us, being manipulated toward my goals, is better than being manipulated away from them.

Absolutely. The positive potential is enormous. The most important thing to realize about these A.I.s is that they are agents, not tools. All previous technologies in history were tools, not agents. An atomic bomb is not an agent. An atomic bomb cannot change in ways that you don't predict. An atomic bomb cannot decide who to bomb. A.I. can.

On the one hand, this makes A.I. much more useful than any previous technology. Because you can be in a relationship with it, and you can tell it what you want, and then it can invent new things that you would not think about. So this is extremely useful.

But the problem is that it's unpredictable and uncontrollable. Do you think you can trust them to just keep to the goals that you're telling them to pursue and not to develop their own goals?

The way that I often like to think about the A.I. revolution at this moment is in terms of immigration. We are about to be or already are in the middle of a major new immigration wave coming to all the countries of the world. The immigrants are not human beings without a visa coming in some boat. They are A.I. entities coming at the speed of light.

Usually, the people who oppose immigration, their main concerns are that the immigrants will take jobs, the immigrants will change the culture, and the immigrants might not be politically loyal.

I'm not sure if that is always true of human beings, human immigrants — but it's definitely true of A.I. immigrants. The A.I. immigrants will take a lot of jobs. The A.I. immigrants will completely change the culture, even things like romantic relationships.

There are people who say: I don't want my daughter to date an immigrant boyfriend. OK, do you want your daughter to date an A.I. boyfriend instead?

And finally, politically, the A.I.s will not necessarily be politically loyal to your country, to your government. At the very least, the A.I.s will be loyal to just two countries in the world, which is the U.S. and China. Down the road, they probably won't be loyal to even those two governments but to themselves.

So should we close the border?

It's interesting. You already see a split within the Republican Party and within MAGA about this question exactly. There are a lot of people there who are extremely concerned and want to close the border. Now, it will not be possible to simply stop the development of A.I.

The question is, as with immigration: How do we build a hybrid society? Because it will be a hybrid society. Society will be a human-A.I. society.

You will have A.I. bankers and teachers and soldiers and border guards. Countries will rely on A.I. border guards to keep the human immigrants away. And A.I. boyfriends and girlfriends and so forth.

The question is: Can we build a good, beneficial hybrid society or not? It will be much, much more difficult than dealing with a human immigration wave because these are a different species. They are not even organic.

I think there are two interesting things that analogy — which is very provocative — pushes you toward.

One is, when you think about how to build a good society around immigration, the thing you're often considering is assimilation. How do you merge the cultures of the people who are coming with the culture that they are coming into? How do you maintain cohesion in that national story that we were talking about earlier?

Do you do that by getting them to learn the language? By more carefully choosing who comes? How do you build structures of assimilation and coherence?

And the other, which is related but different, is in this case they are being pulled in by the government.

When human immigrants come here, it is because they want to be here for a particular reason, right? They are truly agentic. They are here because they want a

better life for their families, a better life for themselves, to have opportunities or freedoms they don't have where they're from.

And in this case, it is the most powerful people in society at different levels who are pulling and accelerating this immigration wave.

Yes.

Some for reasons of profit, some for reasons because they're excited to bring a new kind of intelligence into the world, and at the political level because they want to make sure they get there before China and that America has that power before China has that power.

Exactly.

So what do those similarities or differences to the question of immigration imply for you about what it means to create a structure in which this hybrid society can be healthy?

It's interesting that some of the people who are most vehemently against human immigration are exactly the people who try to force other countries to open their borders to the A.I. immigrants.

This is going to be the major issue of sovereignty for countries all over the world, especially if almost all the A.I. immigrants are either American or Chinese and, down the road, not loyal even to the U.S. or to China but to something else.

One way to do it is to have a ban on A.I. personhood. It doesn't mean to stop the technological development of A.I. It's more of a legal and political issue. Does human society recognize A.I.s as persons?

Now, persons are different from human beings, from entities with bodies and minds. But in many legal systems, like in the U.S., something can be a person even if it's not human.

The best example we have so far are corporations. According to U.S. law, Google is a person. Microsoft is a person. X is a person. Because corporations are persons. And this gives the corporation rights, like owning a bank account, lobbying politicians, donating money to politicians.

It will be extremely dangerous at this point for any country to recognize A.I.s as persons — to allow A.I.s, for instance, to open a bank account or manage a company by themselves.

Previously, when corporations were recognized as persons, this was a legal fiction because all the decisions of the corporation were ultimately made by some human being.

Microsoft is a person according to U.S. law, but every decision Microsoft makes to buy another company, to fire somebody, to hire somebody, there is a human being who really makes this decision. There is no Microsoft who makes the decision.

With A.I., for the first time in history, we have a practical potential for companies without humans. You can have millions, even billions, of A.I.s opening their own companies, their own bank accounts, even hiring people to work for them, deciding on their investment strategy and whatever.

They will have huge advantages over human companies. For instance, the A.I. C.E.O. never sleeps. The A.I. C.E.O. never goes on vacation.

I can imagine some countries — say, a country like Qatar, which has a lot of money, a lot of energy, and very few citizens, saying: Oh, wonderful. I can now have millions of A.I. citizens paying taxes and building companies that trade and do business all over the world. So even if your country doesn't allow A.I.s to build their own companies, what do you do about the Qatari A.I. companies?

The moment you recognize A.I.s as legal persons, this is the moment you really lose control. Because then they can start doing a lot of things in the economic and social and political arena without any human accountability. For instance, donating money to politicians in exchange for the politicians taking care of giving more rights to A.I. persons.

I think that's very, very interesting. I guess one question about whether you call it personhood or not, one of the ways and reasons we think about corporations as persons, which is linguistically a weird thing, is actually to create accountability, to say that the corporation is accountable for what it does.

And one of the fights around questions of A.I. is a question of liability and who is responsible for what the A.I. does.

If you treat them as a person, you could say A.I.s in this world have some liability for what they do — they can be shut down, can be penalized and funded.

There's another question of whether the companies that create them should have the liability. Maybe the people ordering them should have the liability.

But accountability, I think, is downstream actually of liability.

Yes.

Deciding who is punished, who is accountable, for whether that Qatari A.I. company you're talking about defrauds their customers or brings in investment and embezzles it. Who do you sue?

And the companies that produce the A.I.s have a vested interest in not having any liability.

So they are pushing very, very hard for A.I. personhood. They don't want a bill in Congress saying: We recognize A.I.s as persons. There will be a huge public outcry and resistance. They try to establish facts on the ground.

They already succeeded, for instance, in social media. In the universe of social media, A.I.s are already persons. If you have bots creating and spreading lies on social media, effectively, there is almost no liability.

On social media, A.I.s are already functionally persons. If you communicate with someone online, and you think it's a person, but it's an A.I. — nobody is liable for that.

Many of the companies would like to extend this situation to the financial system, to the political system, because it releases them of accountability and liability.

We need to be proactive and have a law that clearly states no A.I. persons. I imagine there would be partisan support for that law, and it will put the companies in a very hard spot because if they try to lobby against the law, they will have to explain to the public why they think it's a good idea that A.I. will be persons. And if you don't think that, why do you oppose the law?

Let me ask you about one other dimension of this here, which brings us full circle, which is the role A.I. is going to have on the stories we tell and the stories we believe. We talked about the way social media and algorithmic media are technologies of fracture as opposed to technologies of cohesion.

I don't even know what stories somebody is getting on their TikTok feed, even if I'm using TikTok sitting in the same home as them. Our ability to even see what we are disagreeing about, to know the sources of those disagreements, is weaker than it has been at any other time.

There's been a lot of discussion and some research on the way that A.I., so far, seems to be something of a centralizing technology. The different models tend to converge around similar answers. They are trained on similar corpora of data. They all seem to be somewhat liberal — liberal in the sort of philosophical sense that we were describing earlier.

You see this on X when people are asking Grok, which is not my favorite A.I., to fact-check things, their ability to help people correct information.

You were saying earlier that we've gone from trusting people to trusting algorithms. Yes.

The algorithms we trust are very impersonal and faceless right now. We don't have a relationship to them, but you're watching people move to trusting A.I. algorithms. And maybe that's better than what they've been doing. Maybe that is more likely, in most cases, to give people a reasonable answer for a question than searching for it on Google or YouTube.

Is there some possibility — and would it be good or bad if there's this possibility — that A.I. is a homogenizing technology? It is a technology that pulls people back toward not a single set of answers, because different people's A.I.s respond to them differently, but generalized in a way toward consensus answers, which every A.I. model we know of seems to prefer when it is done training.

I think that there is a chance, though it's not a certainty, but there is a chance because, in the training of A.I., there is a very high cost to disregarding truth.

To take a concrete example, let's say that you're Russia, and you're trying to develop your own Russian A.I. You give it access to an enormous amount of data and information. Otherwise, you can't train your A.I.

But when somebody asks in Russia or outside Russia: Is Russia a democracy? Is there freedom of speech in Russia? — you want the A.I. to say: Yes, of course. The Russian constitution guarantees freedom of speech, and Russia is a democracy.

This will mean that you need to explain to the A.I. why it needs to lie. And how do you train an A.I. to lie only in certain cases and not in all cases? That's a very difficult engineering challenge, which people did not have with the social media algorithms.

There's good evidence that when you do it, it degrades the overall performance of the A.I., which I found to be a very fascinating thing.

People have tried to do this, and it creates very strange downstream consequences. Like when Elon Musk seemed to give a directive to xAI to make the A.I. less woke, and all of a sudden, it was talking about white genocide everywhere.

It's not easy to turn the dial ideologically and just get a pinpoint outcome of that.

Exactly. If you tell the A.I. that you are the government of Uganda, and you think that there are no gay people in Uganda, all the gay people in Uganda are just brainwashed by Western propaganda, and you want the A.I. to give this answer, the A.I. will need to ignore a lot of scientific research on human sexuality and on what causes people to have this or that sexual orientation.

How do you explain to the A.I. that you need to ignore articles appearing in scientific literature in this case but you can trust them in other cases? It's a very difficult engineering problem.

If this is the top priority of the regime — like if you are Saudi Arabia, and you have billions and billions of dollars, and you want to make sure that the A.I. will not criticize Mohammed bin Salman, you can do that if that's your top priority.

But you can do that only with a few cases. If you try to do it with too many things, you will get a very crappy A.I.

So does this on some level make you optimistic? Because something I've seen you say in different pieces and interviews is that the most important thing is for countries, societies, institutions to have mechanisms of self-correction.

And often the way we build mechanisms of self-correction is to not rely on individual humans being able to aggregate information at that speed, but we have things that are vast, impersonal, not even fully understood. Like markets, where prices flow through very quickly.

It's not that the market cannot fail — it fails all the time. But as a mechanism of self-correction, it is able to move information through very, very rapidly, and it's quite good.

One way I think modernity has been somewhat troubled is that it is much more complex than most of our mechanisms of self-correction can keep up with.

There's just more information than humans and institutions can absorb. Arguably, A.I.s in this telling are additive to our powers of self-correction.

They are an ability for us to have an agent traversing the world on our behalf, institutionally and individually, that is somewhat truth-seeking, at least in most of the cases so far that we've seen. And that gives us the ability to navigate a more complex modernity with a few more resources at our disposal.

When I'm trying to be optimistic about it, this is the form of story I somewhat believe. I'm curious how you think about it.

Information does two very different things in the universe.

Sometimes you try to analyze information to discover something about the world — like you want to discover the laws of physics or you want to understand the cause of some disease. In those cases, A.I. will probably be a force for immense good. A lot of the mysteries of the universe, which are beyond human capacity, A.I. will be able to solve for us.

But if people think that A.I. will thereby make the universe more understandable and more controllable, they are completely mistaken. They don't take into account the other thing that information does, which is to create new stuff.

Information doesn't just tell us things about the world. It creates entirely new things. DNA doesn't tell us the truth about the world. It creates new things, living beings, living entities.

A.I. will tell us the truth about many things. But it will also create a lot of extremely complicated systems that will be far beyond the human ability to understand and control.

These systems will probably dominate our lives, and we will find ourselves not being able to understand our lives anymore.

Maybe the best example, again, is markets, finance. If you think about the financial system, money is the greatest story ever told. It's the only story that almost everybody believes. It's a story in the sense that it's not an objective reality.

The U.S. dollar is just a story we all believe. It doesn't come from the laws of physics. It doesn't tell us something about the universe. We tell the story of the dollar, and as long as everybody believes in it, we can take a dollar, give it to a complete stranger and get bread in exchange.

A.I. will not tell us the truth about finance. A.I. will create an entirely new financial system, which is orders of magnitude more complicated than the one that we have created and that humans will be utterly incapable of understanding.

We will be like horses in the market. When you trade a horse, the horse can see that something is happening in the physical world. The horse can see that I'm giving you the horse, and you're giving me this shiny metal disc.

But the horse doesn't understand what money is. What is this shiny metal thing? Why is it important? You can't eat it. You can't drink it. What is it? We understand; therefore, we control the world and not the horses.

A.I. will create a new financial system that we will not be able to understand. We will see things happening, like: This company fired me, that company hired me. Why? I have no idea. The A.I. has just made some financial transaction that is just orders of magnitude beyond what my mind is capable of understanding.

The history of finance is that over time, people invent more and more sophisticated financial devices. You have coins and then bank notes and checks and bonds and

stocks and E.T.F.s, exchange-traded funds, and C.D.O.s, collateralized debt obligations.

The C.D.O.s were invented by a tiny number of investment wizards and ingenious mathematicians. Almost nobody understood them — certainly not the politicians who were supposed to regulate them.

For a few years, everything seemed wonderful. People were making billions of dollars because of these C.D.O.s. And then the system crashed.

It is very likely that we will see the same thing with A.I.s on a much larger scale.

The same way that we've already seen A.I. invent new ways to play chess, they will invent new ways to invest, which may be much better than what we can come up with. So they will gain more and more power in the financial system, and it will become so complicated that the number of people who understand finance will go down to zero.

And what does it mean for democracy, or also for dictatorship, when nobody — not the president of the U.S., not the president of China, not the president of Russia, not the chiefs of the central banks, no human being — understands finance anymore?

This will be a very big challenge in the coming decades.

That brings up two things for me that I think are worth thinking about. One, Timothy Lee, who writes a great Substack called Understanding A.I., had this piece on why he doesn't think the A.I. scientists are going to work out the way we think they will.

The thing he notes is that we're already seeing examples where A.I. can solve a problem — but not explain to us, in a way that appears to be true, how it solved it.

Not that it's being deceptive. It's just its capacity to pursue the goal and its capacity to explain or even understand how it pursued the goal are not connected to each other. So it's functionally confabulating an explanation for what it did, and then you look into it, and that's not what happened. It did get the right answer, but we don't know how.

So we actually can't learn from it. That's one interesting dimension where you could have these forward leaps in science and other things, but actually, the human stock

of knowledge is getting better at a much slower rate than the number of answers we're getting because we are not learning from the process the way we do when a scientist finds a new answer.

Maybe the counterargument to that is to say that this is perhaps already true about human society in ways we don't always admit. Markets are an example. People often used to say markets are doing things, acting. They don't have agency, but they are a complex, information-using process that leads to outcomes.

And the market cannot explain what happened. We have principles, but often markets act in ways that defy our expectations.

It is already the case that our world is built on systems, organizations, institutions, that they're not like us, they're not conscious, they cannot explain themselves, but they are structuring the world around us. A.I. is more like a market in that way than it is like an entity.

That's absolutely true. The only caveat is that until now, humans were always a limit on markets, on nations, on the financial system. Ultimately, you needed humans to understand something to make the decisions, because nothing else could make the decision.

So A.I. allows all these structures that we've built for thousands of years, which have become more and more complex. A.I. now allows them potentially to cut the connection to humanity and go on a trajectory that is far beyond what the human mind is capable of understanding.

It even happens, in a way, with language itself. The most important inventions or creation of humanity ever until now was language, because it's the basis for everything. Mythology, finance, nations, religions: They are ultimately based on language. Language is essentially glue. It connects things. It connected human beings for tens of thousands of years.

Now, as it frees itself from human beings, it can start connecting in ways that are way beyond our imagination.

In many ways, A.I. is language liberating itself, releasing itself from the control of human beings and starting to explore all the things that language can do when it's not tied to these packages of meat walking around on planet Earth.

It's not consciousness. We talked about it a bit earlier. When the A.I. says, "I love you," does it really feel anything?

One of the biggest discussions in human philosophy for thousands of years was: What is the relationship between language and feelings? The reality beyond the language? This discussion will become, I think, maybe the most important discussion in the world. Because suddenly, what we couldn't imagine for thousands of years is happening: Language is getting out of our control and starting to just do things in the world.

I think that is a good place to end. Speaking of language, what are three books you'd recommend to the audience?

One book about A.I. that I would recommend is Benjamín Labatut's "The Maniac," which is a fictionalized biography of John von Neumann, but also a very imaginative and powerful exploration of the origins of the A.I. revolution and of the potential consequences of it.

Another recommendation is basically any book by Frans de Waal. I really like his first book, "Chimpanzee Politics." I read it 20 years ago, and it completely changed my understanding, not so much of chimpanzees but of human beings and of politics.

I would recommend Stephen Miller, for instance, to read "Chimpanzee Politics." Because the main message there is that politics is not just about force. If you think you can become the alpha male of the chimpanzee band by going around and just beating everybody, you will not survive long to learn from your mistake.

Another book that I would like to recommend is Aldous Huxley's "Brave New World," which I think is maybe the best science-fiction book of the 20th century — certainly the most prophetic.

He wrote it in the 1930s, against the backdrop of the rise of fascism and Communism and so forth. He foresaw that maybe the most effective way, and even the most dangerous way, to control human beings is not by sheer brute force and fear and

terror, like in Orwell's "1984." But if you work with the pleasure principle and with human greed and desire, you can get further than if you just try to crush people and terrorize them all the time.

Yuval Noah Harari, thank you very much.

Thank you.