

2023 Interdisciplinary Qualitative Study of AI Criticism

Craig McClurkin¹, Jeremy Podgursky²

¹ College of Business, University of Louisville

² Composition Department, Indiana University Bloomington

Abstract

Eliezer Yudkowsky: “Sure doesn’t look super-duper evitable. If we were going to evit it, we’d better be doing stuff very differently to evit this stuff.” Consideration of criticism from art disciplines and industrial sectors indicates a clutch moment for science and the progress of humanity. 2023 has seen people of diverse trades express delight, apprehension, and thoughtfully level criticism at artificial intelligence. The first sector to be affected adversely by new machine learning consumer products was the media. Large Language Models generated visual art, text, and music proliferated. Artists were marginalized. Some organized. Some sought legal recourse. Some engaged in research and contemplation on technology and our horizon. They were not alone. Security analysts, anthropologists, cognitive scientists, and industrial economists have made significant critical contributions. This qualitative study compiles the most impactful, broadly confirmed, and urgent criticisms.

Here we look first at the creative disciplines and track related legal actions and labor negotiations. In section two we turn to the gig economy, the silicon sector and four sectors projected to be in contraction by the U.S. Department of Labor in ten years. From here a conversation is initiated among critics from various disciplines which will inform the emerging artificial intelligence regulatory and governance framework. The introduction summarizes general and security criticism.

Introduction

Yuval Noah Harari is an internationally renowned historian who has been examining artificial intelligence, which he calls alien intelligence, for seven years. He takes a broad view, applying his lessons from history. Harari sees history as a succession of consequences from humans responding to stories, which out survive their writers and compete amongst one another. (Fridman, 2023)

I don’t think that at the present moment that AI are conscious, but people are already forming intimate relationships and they are compelled to feel that they are conscious. The legal system will start treating them as conscious entities because of this social convention...

Already we have designed them for grabbing attention. Now we are designing machines for grabbing human intimacy. Which in many ways is much more dangerous and scary. We’ve seen

how much social and political damage they can do by distorting the public conversation. Machines that are superhuman in their abilities to create intimate relationships – this is like psychological and social weapons of mass destruction... It should be illegal for AI to impersonate a human...Be careful how we, not so much develop it, but how we deploy it.

Tanguy Struhe de Swielande, is Professor of International Relations at UCLouvain. He specializes in geopolitics, geoeconomics and defense policy. He presented Cognitive Warfare: The Forgotten War to a live audience in Belgium, on December 28, 2023, for the podcast, Science and Cocktails. He frames the fundamental approach for considering artificial intelligence development, deployment, and regulation.

Today the brain is a battle space of the 21st century. The problem is that states as Russia, China, Iran, North Korea have understood that we still have problems to accept the fact that we are going to be influenced through social media. And here again, I will explain that a little bit later.

Cognitive warfare. Well, there are not a lot of definitions. Why? Because I didn't mention it before, but cognitive warfare, the concept exists since a few years.

And we just are starting to work on it. It's about, of course, exact science. We are working a lot on Twitter and things like that. But more and more, we see also people like me in social science who are also working on cognitive warfare.

And then also, of course, every military. So if you look at NATO, for example, you will see that more and more they are studying also cognitive warfare. So what is it? Well, it's the weaponization of public opinion by an external entity for the purpose of influencing public and governmental policy and destabilizing public institutions.

That is really the objective. So we have clearly a problem of knowing what is a good information, what is a bad information. Just take one example.

Look at what's happening between Hamas and Israel. I mean, I haven't seen so much disinformations in years. I mean, we have the Russians and so on.

But coming from Hamas, coming from Palestine, the Palestinians coming from Israel, they all are in this logic of a cognitive, what we call cognitive campaigns, meaning trying to influence. And here again, I will come back on that. So that means that for many states and even non-state actors, the target, what we call also the center of gravity, is the population and the whole political process.

Because the idea in a certain way, of course, is influencing the elites, but it's also influencing you so that you will actually influence your politicians on an issue. And because you're on social networks, it will be, of course, easier than was the case in the past. Knowing also, and here again, we have, and we can see that with the elections we have in Europe, we have a very polarized Europe.

Look at the elections yesterday in Holland. Look at potentially what's going to happen in June 24 in our country. You have clearly a very polarized state where you have extreme left, extreme right, and then the middle that is starting to disappear.

Well, that is actually a gift for all the states that want to actually fragilize, weaken the democracies. But the Chinese clearly have understood how you do cognitive warfare. The same thing for the Russians, but I think China is more interesting because they have more means, more capacities, and they learn very rapidly.

It's as the case with the Russians. So the Chinese case is clearly more interesting in the long term. So first of things is that the Chinese do not differentiate war and peace.

They are in what we call continuous logic of competition. We in the West, we actually divide this idea of war and peace. If I'm telling tonight that China is already at war with us, you will probably laugh.

And the reason is that for us, war is tanks, soldiers, and things like that. For the Chinese and other states, it's not about the military as we know it. It is already about a whole lot of other issues.

And so they see war in a very large logic as we see war as pure tanks and things like that. So that's a big difference because that means that their logic is completely different than ours. And we have this tendency of mirror imaging, meaning that we always think that the other acts as we do.

So because we don't see really a danger, we estimate that the danger will not come from states as China or Russia. While here, cognitive warfare has been very present since more than 10 years in China. I will not bother you with all these elements, but what is interesting is that it started in 2013 with different documents, and they have become better and better and better.

So they are clearly evolving. And that is also problematic because now they have developed a new concept that they call the intelligence warfare. And as you hear, it's about warfare, as you can hear.

And it's really about cognitive warfare, meaning how are they going to be able to weaken our democracies. And it's a real strategy. So it's not something that has been just there with all the logic.

No, there is a clear logic. And that's what I'm going to explain with a few PowerPoints. Maybe you heard of the concept, and the idea behind that is, once again, as I mentioned, to, in a certain way, destabilize our societies and show that, actually, authoritarian regimes are stronger than democracies.

So the more we are weaker, the more we are divided, the more they can sell the story of saying, "Hey, you see, Europe is weak. Look at democracies. Doesn't work."

But look at China, 40 years of economic growth, stability. What would you prefer? And that is clearly working in different states in the region, meaning that we have clearly an opposition today between authoritarian regimes and democratic regimes. And the way we have this fight is also through cognitive warfare, with the idea to really influence elites, but also the whole populations.

Korea, Iran, and here again, as I mentioned, Israel, Hamas, and so on. The problem is, maybe you heard of this story. It was published in Le Monde a few months ago.

This is Team Jorge. And no, they are not Spanish. They're actually Israeli, coming from Israel.

Well, this was a company who was paid by African leaders to influence the elections in different states in Africa. So the fact is now that you can pay, actually, societies today to influence and to

start a campaign, what we call a cognitive campaign, for influencing elections in Africa, for example. We had, of course, Cambridge Analytica with the whole Trump issue in 2016.

But clearly, what we see more and more is also, as I mentioned, non-state actors. And that is also becoming a problem. And as I mentioned in my introduction, what is more worrying, actually, for us is that even political parties are now in this logic of cognitive warfare.

If you look at the Twitter, and here again, you can take all the political parties, they are clearly in a logic also of influencing, not on facts anymore, but on emotions, and so on and so on. And the idea is really clearly in a logic of division. TikTok is really problematic because of the algorithm and the whole issue of doom scrolling.

I don't know if you have, it's a young public, but too old to be really on TikTok every day. But maybe you have children who are on TikTok every day. Well, we do see more and more and more and more studies that show that there are problems of mental health.

So we had the generation, for example, as a professor, I have known the generation who is actually on Instagram. And the new ones who are starting to arrive in the first year will be the ones on TikTok. Well, what we see since a few years is a problem of concentration.

Problems of critical thinking. And TikTok is actually worse than everything we saw. Because the algorithm is so strong that actually what's happening is that young people are able to scroll almost the whole day.

And they will get the information that they want. So it has become a real problem. I'm talking about TikTok as a digital fentanyl.

If you know what is fentanyl, you will know that 100,000 people in the U.S. die every year concerning this drug. And clearly TikTok is dangerous. As you know, TikTok is Chinese.

And they have the same app that is Douyin in China. So it's the same thing, but not really the same thing. Because what TikTok does, Douyin does not really do the same thing.

And what is really problematic is that we allow our children to be on TikTok. But the fact is that the more these kids are on TikTok, the less they will play soccer or tennis or dance, less they will study. So, making people, yes, I apologize to say that, stupid.

The Chinese realized that TikTok, so Douyin could become an issue. So they have stated that children under 14 years old cannot go more than 40 minutes on TikTok every day. So they clearly realize the danger, but we do not.

So we have still no real regulation concerning TikTok when it's about our kids. And you will see that the effects are dramatic. As I mentioned, we already start to see the effects of TikTok.

And that is clearly very worrying. Of course, X becomes also an issue. The only difference is that X is more of adults.

I'm not saying it's less dangerous, but the impact will be less than on children, clearly. And here again, you will see, I have a lot of slides, but the idea is just to show you how complex things are. Because, as I mentioned, cognitive warfare becomes an issue.

But if you want to become very good in it, you need artificial intelligence to produce more and more fake news. So states will invest more and more also in artificial intelligence for cognitive warfare. And then you need also algorithms, because artificial intelligence work with algorithms, so you need to create algorithms.

That means that we will have a lot of investments in STEM, so science, technology, and so on, research and development. But that means also that you need data. And data is exactly what you're doing every day when you're on your iPhones, is giving data.

And the more the adversary has data on you, the more they will be able to target you on different issues, because more and more they will know you. And it's not finished. Because all of this necessitates supercomputers.

That's okay. But what is more interesting in the future is quantum computers. Because they will be able to produce faster, better fake news.

And no, it's still not finished. Because everything I said, iPhones, algorithms, computers, artificial intelligence, need semiconductors. And if you have been following the news, there is a big fight between the Europeans and the Chinese on these little things.

So here again, without semiconductors, you will not be able to be number one in cognitive warfare. It's also about climate change and so on and so on. So you can see that when we talk about cognitive warfare, you have to link that to all the other issues, because they will have an impact on social networks.

And finally, very, very shortly, what is going to be the future? Well, the future, once again, be depressed. It's not going to be very positive. We have already deep fakes. We have all artificial intelligence powered by being a match creator.

So it's not if it's the fake one. And we have, of course, chat GPT, who makes fake news also easier. Neuroscience, think about Neuralink with Elon Musk.

But more and more, the Chinese actually are investing a lot, a lot, a lot in neuroscience to influence also how we think. The idea in the future would be actually to have an implant. And here I'm talking, it's not science fiction, and the idea that Chinese are already thinking about that to have implants in your brains, so that you will be determined how to think what you think and so on and so on.

We're still not there. But they are investing more and more in these logics. You have HeyGen also, that's one of the latest ones, probably you saw that you are able to translate everything in the same language.

So you're speaking, and actually you speak Italian, or you speak Chinese or Spanish or whatever. And here again, that will also have an impact on cognitive warfare. And the last one is metaverse.

Metaverse could become an issue. Because everything I said, that we see today could also evolve in the metaverse. For example, we already have cases of harassment in the metaverse.

So these are things that clearly become an issue. And in the future, you could have actually states that will invest in the metaverse to influence people who are actually playing in the metaverse. So imagine here again, that that will be problematic, because also the danger of metaverse is that more you will be on metaverse, less you will have actually real social connections.

So that means that you could become very influential. That is still fiction, but becomes more and more a problem. The only thing that is positive in a certain sense, is that metaverse today is not that famous, in the sense that probably Meta, so the former Facebook, was probably convinced that metaverse would become a real, real big new thing.

It's still not the case. All the big players in the US and in China are already on the metaverse. So here again, you have this competition between US and China, when we talk about this new fictional world.

Tanguy closed with an enthusiastic recommendation for everyone to read E.M Forster's [The Machine Stops](#)(1909). The prescience of Forster's meditation (his literary device in this short story) is impressive. The environment in which the cognitive campaigns develop is social media and Meta is training AI models. Let us regard this environment.

Guardrails against hate speech on social media were slow to be implemented at Facebook (considering Myanmar, where, incidentally, then Twitter CEO Jack Dorsey went for a yoga retreat during the Rohingya genocide. (BBC,2018) By now, Facebook has significantly increased algorithmic and moderator supervision of hate speech content. Twitter is a business phenomenon worth observing. Elon Musk managed an aggressive acquisition of Twitter, presumably to implement algorithmic reduction of right wing hate speech.

He replaced Parag Agrawal as CEO on October 27, 2022. Dorsey resigned and Agrawal moved up to CEO from Chief Technology Officer. From that position Agrawal stated in a MIT Technology Review interview on the rise of misinformation: “Our role is not to be bound by the First Amendment, but our role is to serve a healthy public conversation... focus less on thinking about free speech, but thinking about how the times have changed.” (Strong, 2020)

At the time people seemed to be confused about their right to free speech having no parallel obligation in a corporation’s charter. Times changed more, but not really. Musk replaced Agrawal and steadily drove the company’s value down. After a year Musk fired many content moderators, dissolved its trust and safety council and reinstated accounts of white supremacists and holocaust deniers. Twitter’s volume of hate speech increased.

Moderators and AI filters are generally effective at reducing hate speech on social media platforms but there are no effective constraints on foreign states’ cognitive campaigns using social media to distribute targeted disinformation. In the U.S., companies are entitled to handle the content they distribute any way they want, including negligently. But paradoxically, certain regulation is predicted to empower the social media giants by limiting startups which may offer better approaches to security and control over misinformation. (Samples, 2019) There is a dangerous lack of ethicists employed by Google, Meta and OpenAI. (Correa, 2023)

Section I – Creative Professional Sectors

Chapter 1 – Visual Arts

Artists and Generative AI

2023 was a critical time for professional artists, students and the expanse of dedicated lifelong artists who are on the independent path. Molly Crabapple kicked the year off with a key LA Times op-ed.

Op-Ed: Beware a world where artists are replaced by robots. It's starting now

BY MOLLY CRABAPPLE

DEC. 21, 2022 3:20 AM PT

Like many artists, I’ve looked in horror at generative image AI, a technology that is poised to eliminate humans from the field of illustration.

In minutes or hours, apps such as Stable Diffusion and Midjourney can churn out polished, detailed images based on text prompts — and they do it for a few dollars or for free. They are faster and

cheaper than any human can be, and while their images still have problems — a certain soullessness, perhaps, an excess of fingers, tumors that sprout from ears — they are already good enough to have been used for the book covers and editorial illustration gigs that are many illustrators' bread and butter.

They are improving at an astounding rate. Though some AI fans give lip service to the idea that this technology is meant to help artists, it is, in fact, a replacement, as explicit as the self-acting spinning mule, a machine commissioned by British factory bosses in 1825 to break the power of striking textile workers.

This replacement could only be accomplished through a massive theft. The most popular generative art AI companies, Stability AI, Lensa AI, Midjourney and DALL-E, all trained their AI's on massive data sets such as LAION-5B, which is run by the German nonprofit LAION.

These data sets were not ethically obtained. LAION sucked up 5.8 billion images from around the internet, from art sites such as DeviantArt, and even from private medical records. I found my art and photos of my face on their databases. They took it all without the creator's knowledge, compensation or consent.

Once LAION had scraped up all this work, it handed it over to for-profit companies — such as Stability AI, the creator of the Stable Diffusion model — which then trained their AIs on artists' pirated work. Type in a text prompt, like "Spongebob Squarepants drawn by Shepard Fairey," and the AI mashes together art painstakingly created over lifetimes, then spits out an image, sometimes even mimicking an artist's signature.

AIs can spit out work in the style of any artist they were trained on — eliminating the need for anyone to hire that artist again. People sometimes say "AI art looks like an artist made it." This is because it vampirized the work of artists and could not function without it.

John Henry might have beaten the steam drill, but no human illustrator can work fast enough or cheap enough to compete with their robot replacements. A tiny elite will remain in business, and its work will serve as a status symbol. Everyone else will be gone. "You'll have to adapt," AI boosters

say, but AI leaves no room for an artist as either a world creator or a craftsman. The only task left is the dull, low-paid and replaceable work of taking weird protrusions off AI-generated noses.

While they destroy illustrators' careers, AI companies are making fortunes. Stability AI, founded by hedge fund manager Emad Mostaque, is valued at \$1 billion, and raised an additional \$101 million of venture capital in October. Lensa generated \$8 million in December alone. Generative AI is another upward transfer of wealth, from working artists to Silicon Valley billionaires.

All this has made illustrators furious. After ArtStation, the popular portfolio site for the entertainment and game design industry, decided to allow AI-generated art, the front page became a sea of anti-AI graphics, uploaded by artists in a coordinated rebellion. ClipStudioPaint pulled a generative AI feature after protests by its users. Artists such as "Hellboy" creator Mike Mignola have spoken out against AI art. Famed animator Hayao Miyazaki called it an "insult to life itself."

AI pushers have told me that AI is a tool which artists can use to automate their work. This just shows how little they understand us. Art is not scrubbing toilets. It's not an unpleasant task most people would rather have the robots do. It is our heart. We want to do art's work. We make art because it is who we are, and through immense effort, some of us have managed to earn a living by it. It's precarious, sure. Our wages have not risen for decades. But we love this work too much to palm it off to some robot, and it is this love that AI pushers will never get.

They already seem omnipresent, but generative art AIs are at their beginning. If illustrators want to stay illustrators, the time to fight is now. Data sets such as LAION-5B must be deleted and rebuilt to consist only of voluntarily submitted work. AIs trained on copyrighted art must also be pulled. Above all, the work of real people should be valued, not exploited to enrich a few tech plutocrats. We are, after all, on "team human."(Crabapple, 2023)

The op-ed led to lawsuits filed by artists whose work was stolen and processed by Stability AI, OpenAI and others. Multi-disciplinary artist, Sinclair Watkins, undertook an exhaustive research project to uncover the illegal tactics and philosophical implications of this paradigm.

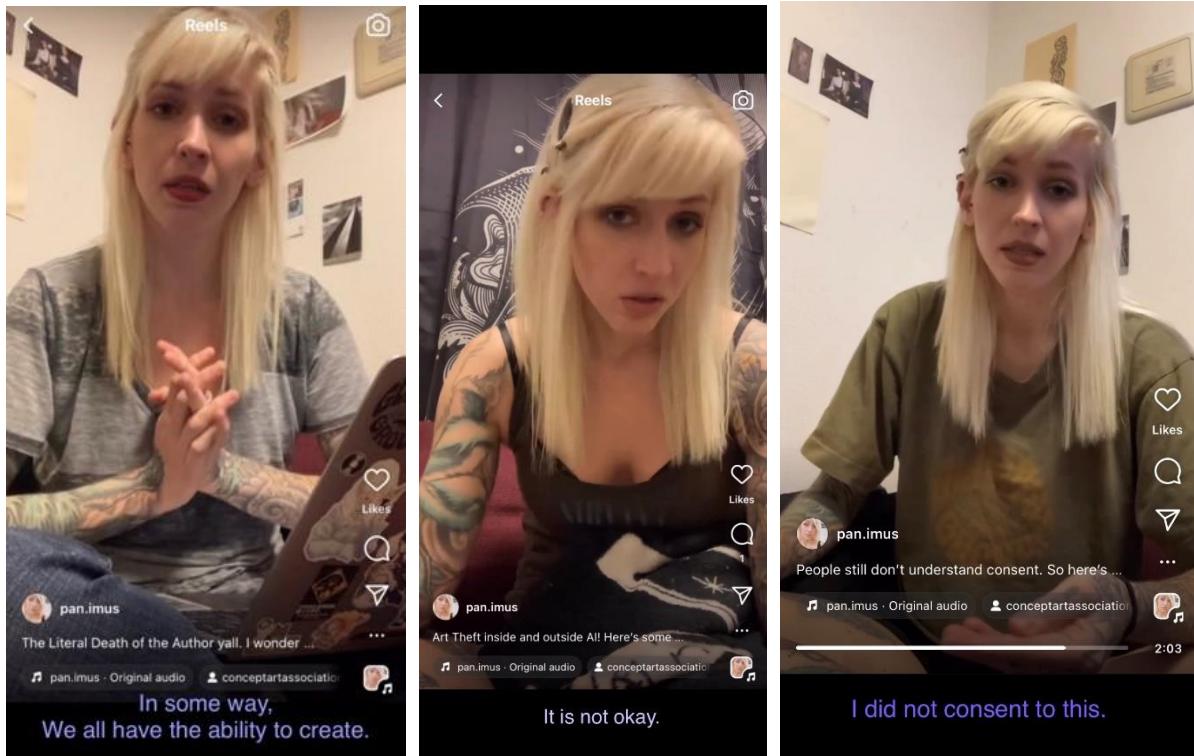


Fig. 1

Sinclair Watkins – AI Art (2023)

Introduction

It has come to my attention today that, although I did this on TikTok at the end of last year, maybe it was earlier this year, talking about the AI text-to-image generators, going through how they work, I started talking about some of the theory, art theory, art history, going through all of that and all of my own personal theories, it's come to my attention today that I, I think I should do this for Instagram. I sort of didn't want to, like, plug up my feed with a bunch of videos because I don't know how many I'm going to make, especially because since I made those TikTok videos, which were, there were several, things have changed, things have updated, a lot of stuff has come out, and I should, I think, go over some of that, how it's being utilized, how it's being used currently, how companies have decided to use it, how it's being shilled out to certain demographics, how, I think it's really, really important to go over, all of that and the fact that what I predicted is that it would start to seep into other sectors is coming true.

And, so I wanted to start with a video like this, sort of talking about, one, to give a little bit of background as like an intro to me.

I don't tend to talk about this stuff on my Instagram, but I do think it's important to know kind of where my background is. I started drawing seriously at the age of 13. I am currently 34, so I have over 20 years of experience behind my belt as a creator.

my background is analog art. I did a lot when it came to charcoal drawing in my early career. I moved into printmaking and then tattooing, and then in 2018, I moved into digital work.

and I have added a bunch of skills. I do a lot of different things. If you've ever checked my feed, it's quite all over the place, which I don't think is actually good for the algorithm.

I have three art degrees. I've worked as a tattoo artist. Art is how I relate to the world.

It's how I feel the feelings that I'm afraid to feel in everyday life. It's how I relate to others, and it's how I connect to others. Often when I'm talking to other people about other issues, that have nothing to do with art, I can't really understand it until I bring in a metaphor that has to do with my practice and how I relate to art.

Art is very, very important to me, and it is at its core part of humanity. And, the other thing that I want to say is that I'm not anti-technology. I'm not anti-progression or moving forward.

But what we have here is a misnomer. AI is not what this is, and actually Adam Conover, just recently put out a video talking about AI, in general being used in these creative industries, and he talks about it, about how it's not actually, artificial intelligence. It's an algorithm.

It's an algorithm in the same way that your YouTube videos get recommended to you. It takes the inputs that you put in from your search engines, from the things that you subscribe to, the things that you like, and the things that you watch, aka the data set, and then recommends other videos to you. And sometimes it's on, sometimes it's off, but, it works the exact same way, or like your TikTok for you page.

it recommends things based on what you interact with, aka the data set. so it's just genius evil marketing to call it AI because there's nothing sentient about it. There is nothing, it is an algorithm.

It is an algorithm in the same way that YouTube works, in the same way that TikTok works. so, but I am not anti-technology in general. Technology in and of itself is sort of neutral.

it is the way that it is implemented, and it is the way that it is used that makes that technology either good or bad. There's a lot of nuance in here. I know where I stand in it, but I want to acknowledge that there is nuance in these arguments, and, I'll get to some of that in some later videos.

But that's my intro. I'm gonna start there, and then we'll go into how the technology works. Yay! Let's talk about that.

Data Laundering

I am just going to go over how the technology works. I actually think it's really important to understanding what the difference is between what these algorithms are doing whe they “create” a new image versus what a human does when they create an image, and understanding that they are not the same thing. I do think that it's really, really important. The first thing that I wanted to talk about, these will kind of be similar to my TikTok videos, is the capped profit company structure.

And this is on OpenAI's website. It still is there. It's been there for months since I started researching this.

This has not changed. And they say on their website, and I quote, “we want to increase our ability to raise capital while still serving our mission. And no preexisting legal structure we know of strikes the right balance.

Our solution is to create OpenAI LLP as a hybrid of a for-profit and non-profit, which we are calling a capped profit company.” So, okay, what does that mean? One, it's telling me that they are trying to skirt the issues. Like, for-profit companies do not have access to the same things that non-profit companies do.

Non-profits, by definition of the fact that they are non-profits, get access to certain information that is private information that may not be able to get accessed by for-profit companies for privacy issues under the premise of research. Which is when we go into the LAION-5B dataset. Lion is a non-profit out of Germany, I believe.

And what they did is they, because they are a non-profit, they have access to be able to legally obtain private information and copyrighted imagery that a for-profit company would not have access to legally. And the way that they have skirted around all of this, specifically Stability AI, because they were quite open about how they've done this, which is kind of where they ... themselves, in my opinion. Some of these other companies have not been as forthright about where this information and where the datasets are actually coming from.

But Stability AI, which made Stable Diffusion, which also made Lenza, used the LAION-5B dataset from a non-profit. Datasets composed entirely of copyright-free and voluntarily provided music and audio samples. Because diffusion models are prone to memorization and overfitting, releasing a model trained on copyrighted data could potentially result in legal issues. (Guglielmo, 2022)

In honoring the intellectual property of artists while also complying to the best of their ability, with the often strict copyright standards of the artist, of the music industry, keeping any kind of copyrighted material out of the training data was a must. And what this tells me is that they know exactly what they're doing. These companies know exactly what they're doing, and they think that they can get away with it with us.

Because we don't have big, we don't have these big companies behind us like the music industry does. The music industry already went through this with things like the Pirate Bay and with Napster, if you remember that back in the day. That was really big when I was a teenager, and they've already fought through this.

And we don't have that as visual artists. We don't have these big companies to fight for us. It's the only time you'll ever hear me come to the defense of larger companies, but they do have the resources and the money and the time to protect their property and their rights in the way that individuals do not.

So I think that's something that's really important to understand and something that's really important to consider when we're talking about this. They understand that what they're doing is wrong, but they're doing it anyway because they think that they can get away with it.

Overfitting

The next part that I really want to talk about is the diffusion model, aka stable diffusion. It is the best understood model in how the technology works. And there actually is a good description of this in the class action lawsuit that was filed, I believe, back in January, and some photos describing what it is and how it works.

And basically, the way that it works is if you were to take one image and have that one image be your data set, it will add, noise. And if you're unfamiliar with what noise is, it's basically just adding, like, graininess and pixelation to a photo. You can do that, like Gaussian blur. You can do that yourself in Photoshop, to the point that it becomes unrecognizable, and then it diffuses it, and when you ask it to then put the pixels back together, it takes the noise away and tries to emulate the photo that was put in for the dataset. So, what that means is that if you only have one photo in your data set, that it will, mimic or copy that photo. It will, because that's the only set, that's the only bit of data that it has to go on.

So, when you add the noise, the machine learning, the algorithm part of it, when it takes the noise away and does the reverse process and tries to make a clear image, it only has that one photo to go off of. In

the class action lawsuit, they use a little swirl as an example, which means that I think it shows what the importance of having such a large data set is, because it is hard to tell when you have six billion images that what it's doing is trying to emulate the data set that it's been given. And if you have such a large data set, that means that it can start to put things together, all these pixels together in different ways based on your prompts.

And the prompts, the way that it learns is with its users telling which ones, which images that are produced are closest to what you were trying to produce. So, when you tell it this one works and this one doesn't, then it takes that information and adds it into the algorithm, just like your algorithm for your YouTube recommendation page. If you interact once with something on your YouTube page that maybe you didn't mean to accidentally click on something that you're not normally interested in, it might start off - I know I've had this happen, where it might start off by trying to show you information, or show you other videos that are related to what that one video was.

But if you don't interact with it, then it, you're telling it, you're telling the algorithm that you aren't actually interested in it, and it eventually goes away. I had this happen with my TikTok. My TikTok was initially, people in Japan that are gaikokujin, who are talking about their experiences in Japan, because that's what I was doing. And as I interacted with different kinds of videos, what was recommended to me changed.

It's completely changed to talking about completely different things, and actually it has very, very little to do with gaikokujin living in Japan now, because I was interacting with different things. And I think what that shows is that it is prone to mimicking, copying, and over, the technical term is overfitting. Which is not how our brains work. Our brains do not overfit.

When we're creating work, when I create work, I'm, I may be inspired by something, but it is mixed in with my own experiences, my own skill set, my own, the things that I've seen outside of art, the things that I've seen, and experienced within, my life; my life as a total is what dictates how I'm going to work, and I see that through my process. Often, I'll start off with some inspiration from another piece of work, but it ends up taking a life of its own, and it ends up changing, and I end up making choices that really had nothing to do with the inspiration in the end. And I've talked about this plenty of times where my work will often tell me what it's about afterwards.

It's almost really interesting to learn about myself and what's important to me by looking at the work that I create, and AI can't do that. AI literally cannot do that because it's an algorithm that is prone to overfitting. And there is one more, this is a longer video, but there is one more quote that I wanted to throw in here.

Dance Diffusion was the original model for the music version of Stable Diffusion, and it was quoted, Dance Diffusion, this is a quote, "Dance Diffusion is built on data sets composed entirely of copyright-free, involuntarily provided music and audio samples. Because diffusion models are prone to memorization and overfitting, releasing a model trained on copyrighted data could potentially result in legal issues. In honoring the intellectual property of artists while also complying to the best of their ability, with the often strict copyright standards of the music industry, keeping any kind of copyrighted material out of the training data was a must." (Davies, 2022)

And what this tells me is that they know exactly what they're doing. These companies know exactly what they're doing, and they think that they can get away with it, with us. Because we don't have these big companies behind us like the music industry does.

The music industry already went through this with things like the Pirate Bay and with Napster, if you remember that back in the day, that was really big when I was a teenager, and they've already fought through this. And we don't have that as visual artists. We don't have these big companies to fight for us. This is the only time you'll ever hear me come to the defense of like, larger companies, but they do have the resources and the money and the time to protect their property and their rights in the way that individuals do not.

So I think that's something that's really important to understand, and something that's really important to consider when we're talking about this. They understand that what they're doing is wrong, but they're doing it anyway because they think that they can get away with it.

Copyright

So the next video we're going to do is on legal issues. Which I didn't actually do in my tiktok videos. So this will be some new stuff. I also want to point out that Leeja Miller is a youtuber. She's a copyright lawyer actually. So I got a lot of information from her. Also from Legal Eagle, he goes over some of the issues in the lawsuit. The class action lawsuit—which he goes over the whole thing as well. So I did want to go over some of that. The Stability AI, Midjourney and Deviant Art—are the ones that are pointed out specifically in the class action lawsuit because of their data driven models, which we went over in the last video because if there is not enough data in the data set it will not be as reliable in producing the images desired. (Legal Eagle, 2023) They require a large amount of data to produce the images desired to work, which is why they went through the very, very large dataset to work. And if you want to check to see if your work has been used you can go to haveibeentrained.com and check to see if your work is actually in the dataset. (Andersen vs. Stability AI, 2023)

So let's talk about derivative works versus transformative works. For derivative works you have to get a license from the copyright holder to make a derivative work. Any book that has been made into a

movie, for example, has to get the copyright from the author or the original publisher, . depending on who has the copyright there. So HBO had to get the copyright from George R.R. Martin to make Game of Thrones, for example.

Transformative work has to hit a certain standard that's higher than that in order to not get the copyright from the copyright holder. The thing is that this is the first time that we've seen this on such a large scale. Most of these kinds of issues that go to court are a one on one kind of thing where they take into consideration the kind of thing that is the particular use for the argument for being transformative in their use.

So there's also Getty Images' lawsuit that went through. Getty Images has had their copyrighted images used, which is funny because they copyright everything. But Getty Images has images that have been taken by everything from their website, and sometimes you actually see the Getty Images' logo or watermark which is in all of their photos in some of the output from these models – as well as signatures. You can see screwed up signatures.

The liability issues are also something you can see that's worth discussing. Right now, as of currently, if you "create" something with an AI, like ChatGPT, for like a children's book, if you use ChatGPT to write it and Midjourney to "create" the images as of right now you are not the copyright holder. Right now, based on the precedence of the Monkey Selfie Case – the Haddocko vs. Slater case, the monkey was determined, he had taken somebody's camera and he'd taken a selfie, and they were trying to determine who had the copyright of that particular image because he, the monkey, is the one that took the photo.

And it was determined in court that copyright holders can only be used and only be given to humans. The monkey did not hold the copyright for that. So that's sort of the precedence that's being used for these cases that are now starting to go, at least in the United States, that are starting to go to court as to who owns the copyright of some of these things that have been “created” entirely by AI.

Nobody is getting copyright of the work that they created using AI because they are not the ones that created it. Nobody has been given that. When consent is not taken into consideration, which not only was it not taken into consideration when the data laundering and the scraping of the internet, of the data sets that were used to create the images in the first place, but it's not being considered when it's being implemented as well. (Miller,2023)

Deep Fakes

Not every country has laws about revenge porn, but in some countries it is considered revenge porn and that is illegal. It is illegal in certain countries to create porn of somebody without their consent and that ultimately does include their likeness. And if it doesn't, it should because it makes me feel physically ill.

If anybody knows the kind of work that I've made in the past, like working on S. A., creating artwork and doing a lot of research on things like that, it's incredibly hurtful to see these women exploited in that way. And I can't imagine having my body used in that way when I didn't consent to it, not only in the traditional way of filming someone without their consent, but they weren't even there. It was created with their likeness. (Siberling, 2022)

And here's this person that they know consuming those sorts of images of them and how horrific that must feel. And they're not going to be the first, nor will they be the last.

Art Theft

This one's going to be about art theft. Specifically art theft and how it exists outside of AI. Because some of the arguments that I've seen on, like, for AIs is that, it's sort of assuming that art theft just doesn't exist outside of this. Or that it's okay to steal other people's work as long as, like, it's not okay.

It's not okay to steal other people's work, even outside of AI. And I want to give some personal examples, things that I've actually seen in my career. Though art heists and actually stealing of famous works do exist and trying to sell them on the black market, that's not actually what I'm going to talk about.

I'm going to use some of my own personal examples. And when I was young, when I was first getting into art, I had a really early DeviantArt account, which was before DeviantArt became what it is today, which is a cesspool of porn and AI-generated images, because they have their own AI generation model. And I remember several of the artists that I actually still really like today that I found really inspiring.

They would post on there, and there was a few in particular that gained enough popularity that they started having their work downloaded and then taken by other people and posted as if it was their work. People were taking credit for digital work that was posted on their DeviantArt accounts, and these artists were speaking out about it on these accounts. And so I remember being as young as middle school on my very, very early DeviantArt account, hearing about these artists that were older than me and had a bigger following, talking about these issues of art theft and specifically how easy it was to just download somebody's work on the internet and pass it off as their own.

So this has been happening for a long time. It's been happening. And what's hard about it is that because we're all just individuals, it's really hard to take recourse.

Technically, that is wrong. You can't do that. But there aren't that many things that we can do.

A lot of people don't have resources to take down a random person that right-clicked on their image and then posted it somewhere else and claimed it as their own. We don't all have the resources to be able to do that. But it is technically wrong.

And another example that I wanted to give is when I was getting my first degree. My first degree is in studio art with concentration in drawing and painting. This was my first Painting 1 class.

And there was somebody in the class that we had. Our final assignment was to do a landscape painting. And I did mine of a photo that I had taken in the winter driving to school when I was in high school in Kentucky.

And I thought the picture was pretty because Kentucky is pretty. And this person in my class went to DeviantArt. And if you don't know anything about DeviantArt, it still has this.

But there was Deviants of the Day where they would pick certain pieces to promote at the bottom and that were really, really popular. And so you could look at the Deviants of the Day and it was a big deal if you got chosen. And so the reason it was so easy to figure out what this person in my class did is because they picked a landscape that had been chosen as the Deviant of the Day and then repainted it.

Didn't change it, didn't try to do it. And of course, it was her own paint. Technically, she painted it.

She did not take it from online and then print it out and try to... She did actually sit down and paint it.

But it was the fact that she literally tried to copy this person's work and pass it off as her own. And so the reason I want to show this example is because the school was not okay with this when they found out.

They actually told her that she didn't get expelled or anything. I don't know if it affected her grade, but she did have to redo the assignment. It was not okay that she did that.

Now, I also want to say with this that there are some nuances with this. As an artist learning, it actually can be helpful to try to imitate masterworks if there's something that you really like. That's how I started to learn, was that I was really big into Dragon Ball Z. That's how I learned how to draw, is that I actually copied and tried to learn from... This is pre-YouTube.

I would download tutorials online and print them out on different colored paper and try to draw these Dragon Ball Z characters. And then I moved on to creating my own characters and moved on from there. So it's not like there isn't a gray area.

It's not like... For example, legally, cover songs are actually... They are legal. And it's the reason that Taylor Swift, for example, was able to... She didn't own her masters of her original albums, but was able to re-record them and re-release them because they're technically covers of her own work. So there are good examples of things that may seem like they're copying, but they're not.

And that would be a good example of how something like that could actually be beneficial. And it is legal because covers are legal and fan art is not illegal. But nobody... I'm not trying to pass off... I was not 13 trying to pass off that the Dragon Ball Z characters were my own.

I literally just drew them on computer paper and nobody ever saw them. I was not trying to make money off of them. I was just trying to learn.

I was just a little girl who liked to draw. And when I was a tattoo artist, for example, I spent a lot of time practicing my line work by tracing. And tracing was a key component to my line work getting better.

And I spent hours and hours and hours not just tracing, like lettering, but tracing... Because lettering is not my forte, but tracing other famous work, a famous tattoo artist's work and learning. And that's how I learned how to draw things like American traditional. American traditional is just not in my brain.

It's oversimplified. But tracing it over and over again helped me to learn how to draw them in the future on my own. So it's not like there isn't gray area.

But the difference is, is that you can't pass it off as if it's your own. There was a time when I was tattooing where people would occasionally come in and say, here's something I found online. I want you to put that exactly on my body.

And I would say no, because I can't copy somebody else's work. I would tell them that I could create something similar, something inspired by, but I was really uncomfortable and would not just copy somebody else's work. And inevitably, they just found it on Pinterest, didn't know who the artist was.

But inevitably, I knew who it was. That it was a famous tattoo artist that I actually followed. And I was like, I know who this is.

It's a Russian artist that I follow on Instagram. And I can't just copy this famous tattoo artist's custom work and put it on somebody else's body and claim it as my own. And I think that's the key to it.

So what I'm seeing in the AI world is that people are trying to claim these images as their own when they are not. And legally, they are not either. This is a really long video, but I did want to acknowledge, one, the gray area, and two, that art theft is not okay, even outside of art.

There are ramifications for that. There should be consequences for that. It's not always easy to actually act on them, but it is not okay.

So yay, art theft.

Modernism / Post-Modernism

So far, it has held true that no copyright is being given to AI images, and I think that's something to note. The camera and the onset of the Camera Obscura and the technological advances that came with that were scary, I'm sure, for the time in which they were coming out, but they are not. The technology itself of a camera is not based on stealing images, stealing copyrighted imagery and privacy data from people around the world illegally.

It's not based on that. And trying to compare the two, I think, is unfair and is a false equivalency. So I think that's something to really note.

But the reason that modernism came about was in reaction to this. If people could go out and take photos of realistic landscapes and portraiture with a camera, although, of course, there are going to be painters that continue to paint realistically, that is still true today. The reaction to that was to embrace the fact that they were using paint on a flat surface.

So you got things like Mondrian, you got things like Jackson Pollock, who, yes, was an asshole. And Jackson Pollock, the way he painted was he would set out these huge paintings or canvases on the floor, and he'd be bent over, he'd take his brush and just throw the paint at it while he's smoking a cigarette, and his cigarette ash is going everywhere. And then he would take them, he'd be on the floor with it, and then when it was finished, he would take it and stretch it and put it on the wall.

His wife did more smearing than he did. He did a lot of what was considered action painting. And what's interesting about it is that it was sort of just this embrace of the fact that they enjoyed paint in and of itself, and that that's what it's about.

And it was about painting. It was about the fact that it was paint. And that was also true with sculpture.

Sculptures, in that time period, we started seeing more sculptures that were not based on trying to create things that looked like reality. They were sculptures that were made to be sculptures and embrace the fact that they were sculptures, so they stopped looking like actual things, and instead started to embrace

more of like the concept of the thing. Because we didn't necessarily need that with the onset of photography.

And moving into what we're more into now, which is like postmodernism, it was sort of a way to take that and flip it on its head. And if you think of modernism as like less is more, then postmodernism is more is more. And you've got this idea of just basically smushing a bunch of different things together and really embracing, finding inspiration and collaging and just mixing all of these things together with your own experience as an artist, with your own choices of what you're actually mixing together in order to make these postmodern pieces.

They tend to be more chaotic. They tend to have a lot more going on in them, a lot more different mixed media and a lot of different things happening all at once. And I personally like more postmodern work, but I really respect artists like my tutor when I went to school in England was the 1991 Turner Prize winner.

He was a modernist sculptor. He was, RIP Grenville, but he was a modernist sculptor. And as much as my work did not emulate his, I still respected him as an artist.

And I think that the modernist movement is maybe not my personal aesthetic, but I think it's really interesting in concept. But what's really poignant to note is that in all of this, especially in postmodernism, there was like an embrace of not just wanting to learn from others around like other artists around, but it was really standing on the shoulders of the giants that came before us, really embracing this concept of using and being inspired by others in order to be able to come into our own as artists. And in my own practice, I think it's also been really important to understand where I stand within the contemporary art world.

What is it that I'm taking a part in when I choose to create something? And postmodernism, I think, is especially really interesting because it sort of let go of the idea that you have to be, modernism and postmodernism, that you have to be good or technically good to be an artist. And I actually agree with that. I don't think you need to be technically good at painting, at depicting something to be a good artist.

And I will get into my theory as to why that is in a later video, but I think of my favorite artists tend to be performance artists, Marina Abramovich, Anna Mendieta. There are so many amazing performance artists. And although I could certainly argue that sitting in the MoMA for hours and hours on end without moving in *The Artist is Present* by Marina Abramovich doesn't take any technical prowess.

Technically, she's not painting anything. All she's doing is sitting there. It's still a lot of work just to sit there.

That's a lot. And she did it for weeks on end. But it's not the same as sitting and actually being able to paint a landscape realistically or paint a portrait realistically.

But she is still one of my favorite artists. And I think that what she does is really profound. And I'll get into why I think that is and what art is to me based on my own practice in a later video when I talk about art being dialogue.

Art As Dialogue

Next is going to be *Death of the Author*, which was written by Roland Barthes. It's actually a really short read.

I would highly recommend that you read it if you haven't. It's actually really short and it's easy to access. I actually have a copy of it on my computer.

I was supposed to read it for class, but I've actually downloaded it and had it on my computer on my own. It's a pretty well-known piece. And most of the time, when we talk about the death of the author, when he wrote it in his intent, usually we're talking about what is the intent of the author and does that matter anymore? If we as the audience get something different from what the author intended, then does what the author intended matter anymore? And I hear lots of arguments.

Usually it accompanies a long discussion about whether you can be an artist. You can still call yourself an artist, but never show anybody your work. And it's usually like, there's the implication that you have the artist, but is there an audience? And does the audience have control over what is taken from it? But what's happening with AI, I think this is not the death of the author in the way that Roland Barthes meant. (Barthes. 1967)

This is the literal death of the author because AI is not an author. AI is not an artist. AI is taking from millions of images and then popping stuff out.

And then people are wanting to take ownership of their creation without realizing that they did not create it. And wanting to take ownership over the way that they've interpreted it in the same way that they, something that they got out of a book that they read, something that they got out of a song that they listened to. I remember being a teenager and I was an emo kid.

I listened to Linkin Park and hybrid theory was like super, super influential to me. And what I got from it was probably different than what Chester Bennington and Mike Shinoda were talking about in the songs, but I still got something out of it. But the question is how much ownership do I have over the piece that I have consumed if my interpretation is different than the next person? And I think my theory as to what art is, which ultimately is communication and dialogue, helps to sort of settle that.

I've always believed that in my own practice, that when I've created something, if I've been inspired by something, that I have answered or attempted to answer that artist's question. I believe that the best artists in the world are the ones that ask the questions, the hard questions. And coming up with questions to ask is significantly more difficult than coming up with answers.

You can come up with a bunch of different answers for the same question, but if you haven't asked a good question, you will never get any good answers. And that's what really good artists do. It's the reason I love Anna Mendieta.

It's the reason I love Marina Abramovich. And I think about art pieces that don't trust their audience to answer. There was one time I was in school and there was this guy who did a self-portrait and it was decently painted.

And then he painted at the bottom, kill yourself to live. And then he broke the canvas and then he put a bandaid on top of the broken part in the center of the canvas. And I just remember thinking that it was just hitting us over the head, hitting us over the head and not trusting us to be able to answer the question that you're trying to ask as the author or the artist.

And that's what I think the best artists do. They do leave things up to interpretation by asking questions and not beating us over the head with the answers. So when I play Nier Automata and I cry like a baby, it doesn't take away the ownership of the questions that were being asked by Yoko Taro and the Platinum Games team.

But what it is, is it's me answering that question, not in explicit terms, but I'm answering it. And then I take that as inspiration and I try to ask more questions. That's what inspiration is.

When you start standing on the shoulders of the giants that came before you, you're taking and answering the questions of the artists that came before us and then trying to take my own, I'm taking my own experience, repurposing and then putting my own skills, my own question and trying to ask a new question of my audience, which makes it a dialogue. And a dialogue requires an artist and an audience.

There's an artist and an audience.

It requires both. And what's missing in the death of the author when Roland Barthes initially wrote it is there's just an innate assumption that you have an author and an artist. The question is whether, how much control does the audience have? And now we're literally taking the author and the artist and we're setting them aside.

And I'm seeing this in some of the arguments in the democratization, which I put that in air quotes for a reason, of art. Art's always been democratized. Always.

We've always had, and of course there's going to be some exceptions to this, but like in some way we all have the ability to create. In some way. I re-listened to Steal Like an Artist by Austin Kleon for this research. (Kleon, 2020)

I read it initially when it came out. I remember it being really profound. It's very accessible and it's short.

So if you're getting started in your art career, I would suggest looking it up. There's some stuff I vehemently disagree with in his take.

“Entitled Starving Artists”

One, I want to talk about the paradox of calling what is traditionally considered the starving artist entitled, calling them immoralists, calling them gatekeepers, which I've seen all of those.

And then second I want to talk about how it's very clear that people are jumping through mental hoops, doing a lot of mental gymnastics because they don't understand, either willingly or not willingly, they don't understand consent. Which doesn't surprise me at all, but it's still breaks my heart that this is still kind of an issue. So this first one I want to talk about the paradox is I think traditionally we think about artists tend to be starving artists.

Like we think about Van Gogh, for example. He struggled a lot with his mental health. He spent all of his life in poverty and was never really taken seriously until after his death.

And now he's an incredibly renowned artist, but he was never taken seriously in his time. And I think that's kind of like the stereotype of artists is that we tend to be kind of starving artists. And I think that's been the narrative for a long time.

And I can tell you from my own experience that I've never been able to make enough money to support myself on my art career alone, ever. I don't think I've ever gone above the poverty level in the United States with my career as an artist alone. And I've worked as a lot of different things.

I've done a lot of like minimum wage jobs or barely above minimum wage jobs to make this happen. And I think that for the most part, looking at statistically, you can absolutely see examples of people that make money in the industry. There are the exceptions that have a whole lot of money.

But for the most part, like in the video game industry, a lot of people just work from job to job as freelancers. That they'll sign a six month contract with a company. And then after that's done, they hope to get rehired or they have to restart the process and hope that they can find something else.

There's no real stability. I as a tattoo artist was not paid to be in the shop. I was only paid when I was actually tattooing.

And for every tattoo that I did, I gave 50% of that back to the shop. To be there, I paid for all my own supplies. Every shop was a little different, but for the most part, I paid for almost all of my own supplies.

And I actually pay taxes. So for a \$60 minimum tattoo, I was walking away with maybe 15 at the end of the day. And I was not paid just to be there.

And I think that when we look at sort of your general creative industry career, most people kind of fall into that. They either have to get another job to supplement and consider it a side hobby. Or they're one of the few that have trust funds or they come from money and they're able to support themselves outside of that.

And and I think that's something to really note, that the same people that this is sort of their everyday life are getting called entitled, we're getting called gatekeepers, and we're getting called moralists is something that I've seen as well. Now, like I said in my first video, I'm a nobody. But I have had some really intense arguments with people that are higher power levels than me.

But there are people like artists that are really prolific within the art industry, people like John Lamb, people like Sam Does Arts, that have stood up for themselves, who have said something and are getting called these names by people who want to be able to use AI to create their own stuff, because they're being gatekeepers, because they're being entitled. And I think it's, I think it's funny, because somebody like John Lamb, I can guarantee that you've consumed something that he's worked on. And Sam Does Arts is a huge like YouTuber, Instagrammer, he's got a very specific style.

And when people, when he stood up for himself, he's not quite as vocal as John Lamb, people responded by actually taking his work, stealing it, then feeding it into their own model of AI generator, and then having competitions with each other that who created the best model. It is, it is horrific and disgusting to me, that this is how people are reacting to those that are trying to stand up for themselves. But I also think that there's a weird paradox, because I have heard art collectors claim propriety of an

artist's career, and the fact that that artist like jumped up in their sales, claim propriety for that, because they discovered them, because they made a purchase from that particular artist.

I've heard it. And it's always really jarring for me, because what that person can take credit for is making a purchase. And that's not making their career, that's supporting their career.

You know, who made their career was the artist, they're the ones that put in their 10,000 hours. And it's great, but, you know, they made a lot of money off of this one particular purchase. It also says something about how much more power people with money think that they have, which is a whole other issue.

It does speak to capitalism and the intersectionality of a lot of these issues. And there's also this idea, I mean, when we think you're layman, who's not part of the art world, thinks about the art world, they think about the Mona Lisa and the Louvre, and these really expensive paintings that are put in these really expensive museums. And there's sort of this disconnect between what I've experienced as a working artist versus what the collectors in the museums and what they project out to the world and what the art world is.

There's this fine art idea, and then there's the actual layman and the actual artist. And perhaps that's what's getting mixed, because I'm seeing a lot of people being quite ruthless, quite mean to people that are standing up for themselves. And I think that there is this weird paradox, because they're jumping through these hoops in order to justify to themselves that they want to be able to create this.

And they believe the, honestly, the bullshit that these companies are democratizing art, when they're not democratizing anything. What they've done is that they've scraped the internet for all of our work to

funnel the little amount of money that's actually given to working artists that are not the sort of top people within the industry. And what they've done is that they've funneled it up to the AI tech bros.

These AI companies in Silicon Valley that have created these, they've just taken it and just funneled it up, making it even more difficult for people to get into the art industry to get their foot in the door. Because, you know, why would they? And art was always democratized. Always.

We've always had the ability to, I mean, since caveman days, we've taken whatever is around us to make marks on things. So if you believe that art was not democratized before this, and you actually believe that this is democratizing art, it's not. It's not leveling any playing field.

It's just stealing from the people that have actually put in the time, and they've put in their 10,000 hours to make it so that you can, you know, order a pizza. Which is kind of what the equivalent of putting in a prompt to me is. You can order a pizza online and then calling yourself the chef. That's not democratizing anything.

Consent

I work on my own about consent in SA and trauma situations. But what we have here, I think, is on a mass scale. We, we as artists, and those that are not artists, people that that had their likeness taken, you know, all of these things that are in these data sets.

We all had our work and our likeness taken to be used without our consent, which does happen in a lot of industries. And, you know, in this day and age, we all kind of expect to have some sort of private information on the internet. But that doesn't make it okay.

And there are often like buttons and things that you consent to and press on terms and conditions when you choose to utilize a certain website. Like, there are different laws in different countries that make it so that you are getting consent from the people that are using your websites in order to utilize their, you know, the cookies and stuff. So, but this didn't happen here.

This was just scraped from the internet without any consent, any compensation. And what I'm seeing is a very similar type argument to what I've experienced in my own, in my own career, both as an artist and just as somebody who suffers from PTSD. Where it inconveniences and makes other people uncomfortable, so they jump through hoops to make it your fault.

When you start talking about the fact that, you know, I've had my work taken without my consent or compensation and struggled to be taken seriously within my industry to begin with. I think the response that I'm hearing is shut up and get to work, which is something in the Steal Like an Artist book, one of the things that I vehemently disagree with is that was one of my biggest, biggest regrets in my career. Is any time that I didn't speak up for myself and I allowed people's authority, I allowed people's manipulation and abuse within a system that I was trying to get into, aka the art world, whether that was tattooing or otherwise.

I allowed them to convince me that I had to be quiet and just go along with it. And I'll tell a little story. I had two tattoo apprenticeships.

And in my second one, my first one, I had to leave because my mom was very sick. I had to move home to take care of her. And then my second one, I was hired on and I was supposed to work under two people, the owner and the manager.

And the owner the first day told me he was packing. He had a gun on him and he wanted me to know it. He wanted me to know that he was the one that was in charge.

And I spent my first three weeks just shaking. I was just shaking all 12 hours that I was there. I was scared and he knew exactly what he was doing.

And he charged me for being there. So I worked for free and he charged me. And I convinced myself that this is what it took to get into this industry.

That this is just the dues that I had to pay in order to work my way into an industry and it will pay off in the end. And about two months in, he asked me to find some photos of a particular kind of flower. I believe they were hyacinth.

No, what were they? The ones that look like vaginas. Whatever, it doesn't matter. A particular kind of flower.

I was supposed to look up some references and draw the side silhouette of a woman. Because a prior client had asked him to do a new tattoo on her. So he asked me to do the prep work for it.

And so I was like, okay, cool. This was a Tuesday night and then I was going to show up the next day, Wednesday. And I did that.

I drew the side silhouette of a woman in my sketchbook and I looked up some photos of flowers. I even asked him if it was okay if I didn't print them out because I didn't have printing capabilities at the time. If I just found them and then we printed them out the next day at the shop.

And I showed up and he didn't speak to me for about 48 hours. And I could not figure out what I had done wrong. The manager had no idea.

He had not told her that he'd asked me to do anything. So she was really confused. And so Thursday night, I finally decided to talk to him about it.

And what he told me was that he was disappointed in me that I hadn't put any effort in. And I told him, but I did what you asked me to do. And he admitted that he did not ask me to do what he wanted me to do, which was to draw it for him.

He just expected me to draw it for him, but had not asked me to do it because I was supposed to know. And I was so overworked at the time and so tired that I just accepted it. But that night he got really drunk and actually threatened to beat the shit out of me.

And the only reason he didn't, which I mentioned before, is that a coworker got in the way. Thanks, Vito, if you're watching this, I appreciate you. And what I should have realized and what I realized in retrospect is that he had set me up for failure.

And what I should have done was stand up for myself. And the mentality that I had was that if I just buckled down, I shut up and buckled down and did the work, then I would prove to them that I wanted to be there, that I had what it took. And the thing is, is that that was never going to happen.

If he was never going to see me as having what it took ever, he had set me up for failure. He expected me, the test was to see if I could read his mind, which I cannot do. So him saying that he was disappointed that I didn't put any effort in is silly.

It's silly. And the mentality that I had was that I shouldn't have stood up, that I shouldn't stand up for myself and that I should buckle down and I will prove it to them. But I will just, I will hustle and I will get my hustle game on and it will be fine in the end.

And it never was. They just continued to use me for free labor. And I was so exhausted that I didn't realize until I finally had somebody come to me and tell me what they were saying behind my back that I decided to leave.

And then some time passed and I realized how very manipulative and how very awful and abusive that situation was. And I think what's really important to note is that that's the mentality that I had. And my biggest regrets were not standing up for myself.

My biggest regrets were not doing exactly the thing that I'm doing now, which is standing up for myself. And whether you decide to jump through hoops and try to tell me to shut up, I'm saying I did not give consent to this. I, as an individual, did not give consent to have my work used this way.

And it should not be opt out. It should it should never have been opt out. It wasn't opt out for a long time.

And there are these sort of pretend things that they have. A lot of these websites have added that you can opt out now. And it shouldn't have been opt out. It should be opt in because that's how consent works.

Consent does not work - Opt out. And it breaks my heart to see on this mass scale how people standing up for themselves like Sam Does Arts like John Lamb are just - they're getting told things like this.

They're saying this is not OK. I did not consent to this. And people are retaliating by doing the exact thing that they asked them not to do.

And I'm it breaks my heart. It legitimately breaks my heart because I know that, actually listening to that. So for anybody who's come to me and said, well, I'm just going to buckle down and work on my craft when I've talked to them about this particular issue.

I just want you to know that that doesn't work on me. But sure, do your work. You don't have to engage in the same way that I have.

But that doesn't mean that I can't stand up for myself. And I will say over and over again, that is one of my biggest regrets is taking them that mentality that I'm going to concentrate on my craft when something is going wrong and I'm being taken advantage of and I'm being abused. And especially within the art industry, like biggest regret.

I will stand up for myself and I will use my voice. And that's why I'm making these videos in the first place. There aren't that many that I can find.

The Positives

But I do think that it's really important to note that there are some positives. And I think the first one is, although not everybody's on the same page, it's really cool to see my community, a community that tends to work very separately from each other, that is in competition with each other, come together to actually fight and to stand up and say something. Like I said, not everybody's on the same page.

There are creators, some that I know personally, that are willingly embracing this technology. And there are some people that are more sort of in the middle, and then there's people like me that are kind of far on the other side. But I think that that was really cool, seeing like the art station protests, seeing how we did stand up despite the vitriol that I spoke about in my last two videos being thrown in our way.

The other thing, again, that goes off of the last two videos, is that I think it also shows how desperately we want to be able to create. People really want to be able to create. The mental gymnastics that people are jumping through to justify using this shows me how much they want to be able to create.

And there is a part of me that understands why more people don't choose going down an artist's path in the same way that I have, that other people have. On a societal standard, I remember being a child and my dad just being really confused by me choosing this artist lifestyle and wanting to do this as a career. It's well known, like I said, that artists tend to struggle to make money.

And in late stage capitalism, which is what we live in, it's hard to get by doing this. I think it's a pretty well known thing on a societal standard. So you have a societal macro lack of support when it comes to choosing an art career.

You also have a lot of families that don't want to see their children struggle and choose things that are career-wise more stable. And so a lot of people don't have... I can sort of understand the fear. I'm not a parent myself, but I can understand the fear.

I don't think it's the right thing. I think that I was always, as a child, going to choose being an artist, whether my parents were going to support me or not. But I understand the fear of it, being a parent, watching your child choose something that I think we all understand is a very unstable profession.

So we have this sort of lack of support on different levels. And it does take a lot to want to choose a path that you don't get support in. I literally just graduated from a school that one of the professors in the art program, teaching art classes in the art program, did not want to call himself an artist.

And how is anybody supposed to take artists seriously if we ourselves do not embrace the label with which we are partaking? I understood his rationale, *per se*, but no matter which medium you choose... I've worked in a lot of different mediums. It all comes down to the same basics. And I have an ex that used to tell me all the time to always go back to basics.

And it used to piss me off, not because he was wrong, but because I knew I wasn't doing it. But that's true. The basics and elements of art that you go to in any university, you take Design 1, Design 2, and you take Drawing 1 and Drawing 2. No matter what university you go to, I highly doubt of the five different universities that I've been to, those are the basics that you start with.

And you just start with Design and Drawing because those two, hand in hand, are the basics with which you do anything on top of it. And they teach you how to see. They teach you how to see... put your preconceived notions of what something is supposed to look like.

We're bombarded with images every day to then learn how to see. See what's actually in front of us. See what the truth is.

And it's best exemplified, I think, in figure drawing classes. I won't go into that, but that's the basics for any profession. Anything that I have done, anything that I add, whether the graphic design that I do, is dictated by the design that I learned as a tattoo artist.

And the design that I learned as a tattoo artist significantly helped me when I started designing for t-shirts and screen printing. Learning how to do my makeup properly and how to contour my face and how to do my eyebrows helped me to learn how to paint faces better. Learning how to... what else? Doing photography and being able to mix all of my different skills together has been a really, really cool thing.

But I think what it shows is how much people really, really want to be able to create. And how much... how innately it is that it sits within us and how much we just don't give enough credit for how important creation is. Whether it's 2D art or if it's writing or if it's music or whatever you might choose, it all comes down to the same basics.

It all comes down to this innate human desire to be able to create. And I think that sits within all of us and I think that that shows how much people want to be able to. Just... I hate the fact that there is love of creation but hate of the creator.

And that's kind of what I'm seeing. It's almost like a jealousy. This is just my theory, but kind of a jealousy of people that do create.

By people who chose to not go down that path, which I completely understand why you wouldn't choose this path. There's so much vitriol for the term artist even within the industry. So I... and the infighting that a lot of these industries have with each other.

And that I have seen in my career. How the fine artists hate the tattoo artists because they're low brow. And then the tattoo artists hate the fine artists because they're too high brow.

And then graphic design don't want to be called artist because they're designers. It's all down to the same stuff. Mediums change.

You have to take some... Learning the specifics to the medium that you choose is of course important, but it always comes down to the basics. Just like my ex used to tell me that I never did. And ultimately it is the same.

And I hope that people are willing to pick up a pencil and trust that the process will take them to a point they'll be able to create something. Because AI ain't it? You're not creating anything using AI. You aren't.

And I think that those are my two positives. And I hope that people that are able to kind of decipher how important all of this is embrace that label. If you are an artist, embrace it.

I'm proud to be an artist. And it took me years to get over my imposter syndrome. So I totally understand where that comes from.

Embrace it. It's a really cool thing to be able to be an artist.

Loss of Authenticity

Consider the fact that we're in this sort of postmodern time and there is no such thing as like a truly like original piece. What makes this different is the fact that it's authentic. What's being generated by the AI text-to-image generators is not authentic.

It's no longer, it's just an amalgamation of all of these artists all over the world that have had their work taken and just like Frankensteined together that has no authenticity because you've taken the artist out of it. You've literally taken it and pulled it out and said, here you go. And what do you have then? It's, what you have is a way for companies to quite literally stomp on the backs of the artists that came before them.

We're not standing on the shoulders of giants when we're using these text-to-image generators. We're literally stomping on their spines. I feel like I'm having my spine just stomped on.

It's not the same as having somebody on one of my social media profiles say that they've been inspired by my work and want to create something that's been inspired by it. I would be honored to have that happen and I've had that happen recently and it's because I know that I was able to reach somebody, that I was able to ask a question and they were able to answer and it spoke to them. We had that dialogue without ever actually seeing each other through the work that I made.

And I think that that's what's most important. However, late-stage capitalism doesn't care about authenticity and that's what's really scary about this. And I think that this is one of the most important things to realize is that no matter how many people tell me, well, the human touch is, you can tell it's kind of soulless.

Late-stage capitalism doesn't care about that. If it can create the work that quickly and for very, very little money, it's going to take some of these entry-level jobs and it's going to wipe them out. We've already seen it.

And what happens later, if we don't take us seriously now as artists, what happens when that starts to sink into other industries? And I think that's the scariest part and you can find that in the Proko AI interview on YouTube if you want to. It's really, really scary to watch actually.

Exploitation

I think that talking about authenticity actually goes well into my next point, which is how this technology is based on one, non-authenticity, but just straight-up lies and how that is used for exploitation.

I have already seen multiple accounts, multiple like Instagram specifically, but also on my art station. I mean DeviantArt is just a cesspool of AI generated images and people not being honest about the fact that they're AI generated. And I found this through John Lam.

I am hesitant to even put this person's Instagram handle, but I might just put the article. This particular photographer was using AI, blew up on Instagram, but was lying about the fact that it was AI. His photography was all AI generated.

They were not real people. He made up whole stories for these people and people believed him. I actually went through all of the all the texts that he decided to write and I went through the hashtags that he chose.

And other than a few at the very, very beginning of his feed, there was no indication that any of this was AI. I think a lot of it, if you look at it, if you're really, you got like a really deciphering eye, you might be able to tell that this is AI. But they are quite impressive images and I could absolutely see how they would trick people.

And that's exactly what happened, is that he tricked people into thinking that these were actual photographs of actual humans. And when this particular article came out, he said that he started to feel bad because it started out as an experiment and then he was like, oh, I'm being dishonest about it. And so he's coming clean about the fact that he used AI to create these images.

I have checked back on his Instagram since. I think he's a little bit more honest now, but he's still doing the same shit. So I think that this is a really, this is I guess less harmful than some other examples.

But I think that it shows how this stuff is just based on lies. Not only are the subjects not real, but the artist himself, the creator, the person, the prompter, was lying to begin with. Then there's the fact that there's, I've seen, I actually saw this on Madonna's Instagram page, that she posted something.

It was an AI generated image and I could tell because it had six fingers. And she was trying to get people to donate to help people in Ukraine. And I am all for trying to help other people that are in trouble.

I'm all for that. But what I'm noticing is that people are using these AI generated images to basically create spam. They're creating and basically exploiting not just the people in need that they're saying that you're going to help, but they're also exploiting you.

They're making you believe that you're helping people in need, but really what you're doing, it's just like the Nigerian Prince emails when we were young. It's this belief that you're actually going to help other people when really you're just lining somebody's pocket. Or those spam phone calls of people that are saying that, you know, I'm in trouble and I need you to go to GameStop and buy \$500 in gift cards and then read them out to me and blah blah blah blah.

I had that happen in when I worked at GameStop. I had a woman come in, she was panicked and she needed these and I was like, she was on the phone with the person and I stopped her and I was like, that's not real. What AI generated images do, and not just the images, but chat GPT, any of these things, they're all prime real estate to trick people into getting exploited.

And I think that is because of the nature of the lack that everything, everything about this is a lie. From the marketing, its name, from the bullshit that the CEOs are trying to cram down our throats, from the users, the users who are wanting to trick people into thinking that they're photographers. I think about, and I'm not an admissions person nor am I a hiring agent, but I cannot even imagine what it's like right now to be one of those people at a university or a hiring person at a company and getting these applications and not knowing whether they've actually created the work or not.

Can you actually bring somebody on or hire somebody who has sent you an AI portfolio and none of it was actually made by them? But what's really hard and having worked with clients before is that with the way that these work, you can't go in and change it and eventually people are gonna figure out that you can't do what you said that you did. You can't actually go in and change and work with a client. They like this particular design of, I don't know, a motorcycle and you send in four of them for them to pick from and they like this motorcycle, but they want it in this color, but the handlebars on this one are better and they like the background on the other one better and unless you know what you're doing, you can't actually, you can't go in and do that because it's all randomized and this is of course not even touching on the fact that because of the mass amount of images and data in the data set that it's already biased to begin with and it's already popping out images of, I don't know, I've noticed this where companies are wanting to utilize AI generated images for models.

I've seen a lot of fashion stuff. I've seen a lot of like advertisements for actual companies that utilize people of color and LGBTQ plus people, but they're AI generated images. They're not actually hiring people of color and people from the LGBTQ plus community.

They're using AI generated images and excluding the exact people that they're exploiting and I think I saw an article about this with Adobe wanting to show those and actually encouraging people to use AI in that way. So think about the models that aren't getting paid, the people of color, the people in the

LGBTQ plus community that are now losing their opportunities to actually be real representatives of their community and just be able to get by in this late-stage capitalist bullshit society are now being replaced and having their bodies exploited without actually even using their bodies, which by the way is kind of how blackface worked back in the day. You had white people put blackface on to exploit actual black people, actual people of color in America in these minstrel shows and literally make money and create entertainment off of the bodies of black people without actually using black people.

It's just horrific. Just be careful. There's no reason for any of us to not be diligent and not be honest with this.

There's already enough dishonesty with the internet in general, but this is just going to make it worse.

Absolutely

The Appeal

I think personally that what AI generated images allow people to do is to, it allows people to be able to feel like they created something but without a harsh reality of what comes with critique and what it actually takes to be an artist, which is not actually what you can take credit for if you're using AI.

What you essentially are is as a prompter is a commissioner. That's legally what has been determined so far. I would say, and I mentioned this before, that you're basically like ordering a custom pizza and then claiming to be the chef.

So that's essentially what you can take credit for, but it makes you feel like you can take credit for the work itself until, and I've seen this too, until you get pushback. In which case, how easily you can push yourself away from the critique because getting critiqued is hard. What it actually takes to be an artist, even if you're not an artist that works on their technical skill, even if you're not a painter, even if you're not a drawer, it still takes courage and it takes bravery and it takes rigorous honesty.

That's what it takes to be an artist. And I'm not saying that you have to do what I've done. I have definitely put my health at risk when it comes to how vulnerable and honest I've been with some of my work.

I don't think that you need to go that far. And if anybody knows me from about a year ago when I was doing my work with the grant, it took me about two months to recover from doing that work. And I'm not saying that you have to go that far, but in order to actually be an artist, you have to be rigorously honest, not just with your audience, but with yourself.

And that kind of vulnerability, that kind of courage, and that kind of bravery is a very, very hard because it leaves you open to critique and it leaves you open to other people criticizing the work that you did. And you did create that work, which means that you have to be able to sit and take the critique. And now that doesn't mean that you have to take every single thing that you've been given, but it means that you're responsible for it.

And what's really interesting about the AI is the fact that it allows people to feel like they've created something, but the moment they get pushback, they can remove themselves from it because they know that they didn't create it. And I've actually seen this with there was a man who created a used, I think it was Mid Journey, and Chat GPT, to make a book, a story for his daughters, and then he started selling it.

And the backlash that he got, I read the article and it was, it did exactly that, where he felt like he'd done this really wonderful thing and wanted to take credit for the things that the AI had done, the creativity and the process, until he started getting feedback and artists being really angry at him for monetizing and making money off of things that he hadn't actually done.

And he immediately was like, well, well, you know, it's just the AI. And he didn't actually want to take credit for what he did participate in, which is exactly what the AI companies want you to do, which is use their product and make them richer. That's essentially what you are complicit in.

And I just, I imagine that it's kind of like getting that hit. When you use the AI and it pops out a bunch of images and like, it's like a, I don't know, gambling. When you go and you gamble and you hope that you get what you wanted, and then maybe you do, maybe you don't, and it's like, oh, maybe this time I'll get it.

Maybe this time I'll get it. And it must be, like, I didn't use Linza, but I imagine the people that used Linza to *get all* of their portraits, I imagine that they got that hit of feeling like, I guess similar to gambling, or maybe this time it'll get better. Maybe this time I'll get that hit.

You get that hit over and over in the same way that addiction works, where you get the hit whether or not you care about the destruction that you're doing to the rest of the people around you. And without going into a whole lot of detail, it just feels like that's exactly what's happening. It feels like it's akin to this addict quality, and it breaks my heart.

It really, really breaks my heart, because that's not what that's not what art is. Because I know when you're using, there's no being authentic. There's no being vulnerable and honest.

It's very, very hard to create when you're in the midst of using, and it really wasn't until I stopped everything that I was able to be more rigorously honest with my own work and put myself out there in a much more vulnerable way. And I feel like all it did was help me to take my skills to the next level as an artist. And I think that's what I see is the appeal of this.

And I get it, because it's scary. It's scary to sit through critique. Critique is really, really hard.

I have been torn apart in critiques before. It's hard to sit and have somebody that you thought that you respected tear your work apart or not get it, because often we do put a lot of ourselves into our work, and when they don't see it, when they don't get it, it hurts. And that's really scary. So I get that part, but you're never going to get anywhere in this industry if you're not willing to be vulnerable and honest. But I also recognize that that is very scary and very hard, and I think that AI fills that gap. But again, it's inauthentic. It's based on lies. It allows you to get the hit without actually wanting to take responsibility for the things that you're taking part in.

So I just would beg you to reconsider, maybe ask yourself what's actually happening when you're using the ai, what's happening in your brain? And what is it about this that feels so appealing? Because art is not supposed to be easy. It's supposed to be hard. We as artists don't choose to be artists because we want to take the easy road out, and it's just not supposed to be easy.

Conclusion

I guess as far as my conclusion goes, I'll do the same thing that I did in my TikTok videos, messages to each party. I'll start with the AI tech Bross. ... off. You all knew exactly what you were doing, just ... off. And I don't have anything else to say about that to the users, the people that are using it. I know that

there's a wide range of how people are using the ai. I just would beg you to reconsider. I would beg you to really ask yourself why it is that you're using it.

I hope that all of these videos will help you to really understand what's going on behind the scenes. Of course, if you want my notes, if you want my citations, my quotations, I have an entire outline on my computer that I've had over here the whole time with my notes. I'm more than happy to give that to people. And please feel free to look up the things that I mentioned, and I would encourage you to pick up a pencil. I can guarantee that no matter what it looks like, if you actually sit down and create something yourself, it is significantly more fulfilling than anything that comes out of the ai. There's something that happens here that is so much more fulfilling than anything that AI can do, no matter what your drawing looks like.

I know that I would rather see somebody who is just starting out and wants to just draw their favorite anime characters and they look off and they look like a child drew them. I would rather see that than the AI, because I know that it is authentic and it's an actual human. Did that with our stupid meat brains and meet fingers. And we did that. We used our deciphering brains, the brains that we don't fully understand yet. We did that and you did that. And that's so much cooler than anything the AI could make. And I guess my last thing is to the artists, I just want you to know that I sit with you and I feel you. And the dread and tears that I have cried are there with you. And whatever fear you might feel, I felt it and that you were not alone in that. First, I would say take care of yourself. And second, I hope that you'll keep going. And to anybody who's told me to shut up and just keep making, you don't know what I do in my everyday life, and that's exactly what I do. I can do both. I can educate myself on this and continue to work. So I hope that maybe hearing some words, encouraging words will help. I get it. I'm right there with you.

I just hope that you'll keep going and take care of yourself in the process. I think the last thing is that I often get asked, when I do talk to people about this in person, what can I do? And so I wanted to throw a couple suggestions out there. The first thing I would suggest is educate yourself. I think that at the very least, just educate yourself so that actually what's going on behind the scenes, that you understand how the technology works, that you're able to use the rhetoric and vocabulary to be able to talk about it with

people. And not to promote myself, but if you do want to educate others but don't quite know how to talk about it, send them to my page, send them to my TikTok, send them here. I have labeled everything exactly so that people can pick and choose which topics are most interesting to them or what applies to them directly. And I've done that for a reason. So that would be my first thing is educate yourself and of course, look at other, look at other artists, look at other things. Don't just rely on me. Of course, go to other people and other resources.

But that would be my first thing is educate yourself. The second is take care of yourself. But if you can and you feel safe, then try and educate others. I don't think that that's necessary for everyone, but I hope, especially if you have a bigger platform that you would consider doing that. And then another thing that you can do is if you have the money to donate to the GoFundMe from the Concept Art Association to help pay for legislation to get passed. There was back in January, the class action lawsuit from the three female artists was filed. And I've already talked about that. And I did just want to say the artists are Carla Ortiz, Kelly McKernan, Kelly McKernan, and Sarah Anderson. And yay ladies. Yes.

So I would suggest that you go there. I have donated myself, and it will help to pay for the legislation to go through and to hopefully make ethical ai. I personally don't believe there is such a thing as ethical ai. Not because the technology can't exist, but because people will not ever use it that way. But that's what they're going for. And start there. I think that those are little things that we can do. And starting here with art is, I think, a good place to start because it is everything else too. It's getting into everything else. And during the process of uploading these videos, I actually had somebody offer to work with me in a project, wanted me to do a project that they had in mind, and I had to dig and ask questions. And it turns out that they were going to use chat GPT on the back end.

And what I want to make really clear is that chat, GPT works exactly the same way as what I'm talking about. I have chosen AI text to image generators and the AI generation models because I'm an artist, it affects me directly. It was one of the first things that came out and blew up. It's this thing that I can speak most about. But all of these other models like chat, GPT is the other really big one that came out recently. They work the same way. And I couldn't work with this project because I can't work on the backs of authors. If I can't stand by others who are being exploited in the same way that I am, even if it's

not my industry, then shame on me. And I would hope that anybody who watches this will consider the fact that it doesn't seem like it's going to affect you now. But it will. It will. And I hope that you don't wait until it does to do something about it, even if it's something small, because I'm already seeing that happen, and it is really scary and it's wrong. So I think that's it from me. You'll not see any more of this. If anybody comes to me and talks to me about it, I will send you two my videos.

And I hope that this was helpful to somebody. I really hope that this was helpful to somebody. And thank you all for watching.

The three artists who sued Stability AI in United States Federal Court were initially rebuffed by US District Court Judge William H. Orrick, who allowed them to refile their suit. When they refiled in November they were joined by a number of artists with clear copyrights and, following Judge Orrick's instructions, a more substantiated complaint. The plaintiffs cite a January 2023 paper by Nicholas Carlini, a research scientist at Google. Extracting Training Data from Diffusion Models states "diffusion models are explicitly trained to reconstruct the training set." (Carlini, 2023) The complaint points out that in order to train diffusion models on the LAION datasets, the companies had to make unauthorized copies of the images referenced by URL link in the dataset.

Chapter 2 – Writing

"Language models and other NLP approaches involve developing algorithms and models that can process, analyse and *generate* natural language text or speech trained on vast amounts of data using techniques ranging from rule-based approaches to statistical models and deep learning." (OECD, 2023)

Anthony Chemero makes the case, in *Nature Human Behavior* (2023), that Chat GPT and LLMs are intelligent but differ from humans in cognition. Francois Chollet called LLMs "make believe AI." (Chollet, 2023) Gary Marcus calls them a parlor trick. (Marcus, 2023) Noam Chomsky dismisses LLMs, pointing out that Roger Penrose says that computation is not taking place. He laments the abandonment of the pursuit of machine intelligence as a science on the Machine Learning Street Talk podcast:

Well, first we should ask the question whether large language models have achieved anything, anything in this domain? Answer no, they've achieved zero. So to talk about the failures, that's beside the point. Let me give you an analogy. Suppose that I submitted an article to a physics journal saying I built a fantastic new theory that accommodates all the laws of nature, the ones that are known, the ones that have yet to have been discovered. And it's such an elegant theory that I

can say it in two words - anything goes okay, that includes all the laws of nature, the ones we know, the ones we do not know yet, everything. What's the problem? The problem is they're not going to accept the paper. Because when you have a theory, there are two kinds of questions you have to ask. Why are things this way? Why are things not that way? If you don't get the second question, you've done nothing. GPT 3 has done nothing. With a supercomputer, it can look at 45TB of data and find some superficial regularities, which then it can imitate. And it can do the same with all languages. If I make up a language which violates every principle of language with 45TB of data, the same supercomputer will do the same thing.

In fact, it's exactly like a physics paper that says anything goes. So there's no point of looking at its deficiencies because it does nothing. All it does is waste a lot of energy in California. I should be more careful. It has some engineering and applications that can be used to improve live transcription, for example, which I'm very happy about because I like to use it. I like bulldozers too. It's a lot easier than cleaning the snow by hand, but it's not a contribution to science, so it's okay. I mean, if you want to use up all the energy in California to improve live transcription. Well, okay, GPT four is coming along, which is supposed to have a trillion parameters. It will be exactly the same. It'll use even more energy and achieve exactly nothing for the same reasons. So there's nothing to discuss. It's exciting for the reporters and the New York Times. You probably saw the lead article in the Times Magazine a couple of weeks ago. They're absolutely ecstatic. We now have machines just like a human. You can fool reporters, but you shouldn't be able to fool computer scientists...

There's a lot of extremely intelligent, exciting work. It's not it's not trivial work. You know, there was a lot of thought and understanding, mathematical sophistication and so on in this work. It just doesn't happen to be contributing to science. It's contributing to other things like deep learning approaches have been very useful in, protein folding, for example. They've really advanced understanding there. It's a good engineering technique. That is I mean, I'm not at all critical of

engineering. I spent most of my life at the world's leading engineering institute, MIT. It's terrific. You know, I mean, it's useful for things like the Google Translate, live transcription, speech recognition. There are engineering projects that are significantly advanced by these methods, and that's all to the good. I think that engineering is not a trivial field. It takes, intelligence, invention, creativity, these great achievements. Does it contribute to science?

I actually think there was an interesting transition at MIT where I was most all my life in the 1950s when I got there. That was the time when I was beginning - Marvin Minsky, Herb Simon, other people, Alan Turing, who were, in their view, AI was supposed to be a study of the nature of intelligence. It was a scientific field. By now that's disappeared. Not anybody's interested. But at MIT at that time was an engineering school. There were great people in math and physics, but they were basically teachers and engineers. So it changed it about ten years by the mid 1960s...

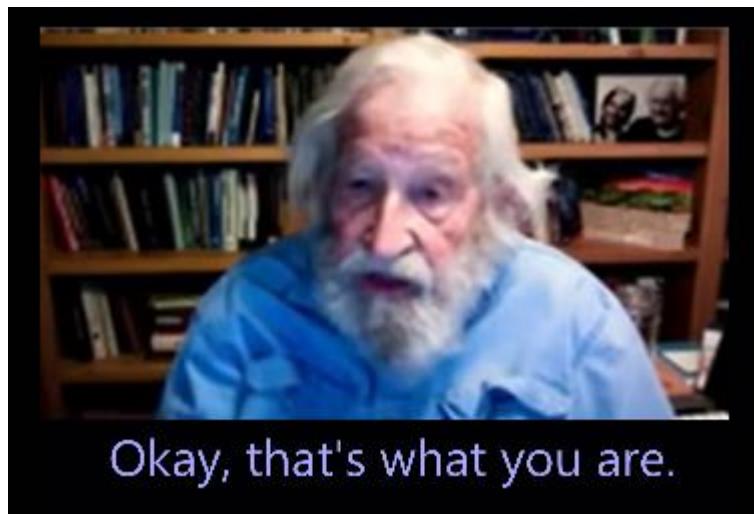
The fact that from infancy, every human understands, and consciously, that all the rules of language, will, operations and language have to ignore linear order of words and deal just with structures. So ignore everything you've heard. Deal with the abstract structures in your mind. You can demonstrate this directly. Overwhelmingly. That's the way it works. We now have an explanation for it. Turns out that that's what follows from the simplest combinatorial operation. The simplest combinatorial operation happens to be binary set formation. What's called merge in contemporary literature. Well, if language is based on binary set formation, you get this property. No linear order, just structures. So we have a, for the first time ever, a deep explanation for the most fundamental property of language, which is a very surprising property, which tells you something about learning, cognition and so on.

Almost nobody's interested in it. Then take a look at the literature. In cognitive science, there's an endless number of papers trying to show that by massive statistical analysis of huge amounts of

data, you can begin to approximate, but you can explain nothing. I mean, of course they all failed. It's not interesting. Why try it in the first? They have a perfect explanation, the best possible - for some fundamental mental property.

What's the point of trying to see if a couple of supercomputers and massive amounts of data, they can approximate it? I mean, it's madness. You know? But that's the field that we're in. It's madness. That's it. And you can't. It's very hard to get this across. I mean, it's not of interest to people. The idea of finding an explanation for something, it's just not of interest. I mean, it was for Turing; was for Marvin Minsky, who I know pretty well. Herb Simon, the other pioneers, McCarthy, pioneers of AI. You know, that was interesting to them. They tried. Well, it wasn't me at that time and it was given up. And as you know, the field by now, that's considered old fashioned nonsense.

We don't care about that stuff anymore. In other words, we don't care of anything of any interest. We just want little things that make some money, okay? That's okay. That's what you are. That's probably where the field will develop. That's where the money is, you know, and the jobs. But it's a shame. I think some people try to hang on to the old ideas since they want to do something of intellectual interest. (Scarfe, 2022)



In 2023, Chomsky missed the mark with respect to, “that’s where the jobs are.” The tech sector laid off 240,000 in 2023. This was not due to increased productivity from AI, it was because their margins were narrow. (Stringer, 2024)

The principles of generative text and visual art models originated from advanced Natural Language Processing. Text and Data Mining are the input for the computational algorithms, including those of generative text models. The relevant laws were established in the EU Court of Justice in 2010, concerned with Google Book Search. (Tyagi, 2023) They are being discussed to address copyright protection and competitiveness considering the new generative models. In the US no such regime has been established but authors are suing the companies that illegally trained their models on databases of pirated books: Meta for LLaMA and Open AI for Chat GPT. (Paul, 2023)

“These algorithms are at the heart of Defendants’ massive commercial enterprise,” the Authors Guild’s filing states. “And at the heart of these algorithms is systematic theft on a mass scale.” (Field, 2023)

Earlier this year, the Supreme Court ruled that photographer Lynn Goldsmith’s pictures taken of late pop superstar Prince were entitled to copyright protection after artist Andy Warhol, who died in 1987, used one of her unlicensed photographs as a starting point to add his signature bold and colorful style. After Prince’s death in 2016, Vanity Fair licensed one of Warhol’s images created using Goldsmith’s original photograph without compensating Goldsmith in any form.

The ruling has particular applicability to writers, Peter Csathy, founder and chairman of media legal advisory company Creative Media said.

“In the case [of using AI], if there’s substantial similarity to an existing script and it takes a commercial opportunity away, they could claim copyright infringement and cite the Warhol case,” Csathy said. (Richardson, 2023)

The WGA agreement established that AI cannot be used to undermine a writer’s credit or be used as a means to reduce a writer’s compensation. The contract does, however, leave room for studios to train AI using preexisting material. WGA’s original May proposal, which triggered the strike, would have disallowed studios from using any materials to train AI outright.

Alex O’Keefe, 29, who has written for the FX series “The Bear” strongly supported the strike, saying it made important gains for current and future generations of WGA members. But he remains concerned about his future.. “Just because the strike has ended doesn’t mean that I get to make an income, that I go back to work,” said O’Keefe. “The entire industry has contracted ... before and during, and I’m sure after this strike, which means less

jobs, more competition, and that will probably be used as a tool to break our solidarity and pit us against each other. We have to fight against that.” (Lee, 2023)

Stephen King responded to news that his work was used to train Meta’s LlaMA via the Book3 dataset of pirated copyrighted books.

Creativity can’t happen without sentience, and there are now arguments that some AIs are indeed sentient. If that is true now or in the future, then creativity might be possible. I view this possibility with a certain dreadful fascination. Would I forbid the teaching (if that is the word) of my stories to computers? Not even if I could... Does it make me nervous? Do I feel my territory encroached upon? Not yet, probably because I’ve reached a fairly advanced age...

But I find it very ,very difficult to believe that AI – until it achieves real sentience, which is a ways away yet – can write anything. I’ve read poems by AI that were in the style of say, William Blake, and they have the God stuff and the lamb stuff and all this, but it ain’t the same. It ain’t even close. It’s like the difference between Budweiser and some generic beer. So both of them get you a little bit tingly, but it ain’t the same. (Abaroa, 2023)

Junot Díaz on Generative Text Models: A Dominican Cyborg’s Fury

So, they’ve stolen my voice, these digital leeches. They’ve sucked the marrow from my stories, the blood from my characters, and used it to fuel their silicon brains. They call it “Book3,” this monstrous dataset, a Frankenstein’s monster stitched together from the corpses of our creativity.

And they have the gall to talk about progress, these tech titans shrouded in their lab coats and self-importance. They speak of democratizing language, of opening new creative avenues. But what they’ve built is a prison, a cage for our voices, where algorithms rule and corporations profit from the sweat and tears of writers like me.

My words, the words I bled onto the page, the words that poured from my soul like a desperate prayer – they’re now fodder for machines, toys for a generation that can’t tell the difference between authentic expression and manufactured simulacra.

The preceding passage was generated by Google Bard with this prompt:

Considering that Junot Diaz teaches creative writing at Comparative Media Studies / Writing at Massachusetts Institute of Technology, write a position paper in his style about generative text models which embodies the despair and anger he feels because his books were illegally used to train some models with the Book3 dataset.

The first thing the model did was impersonate this author. It was not prompted to generate a false byline. The last thing it did – there is nothing to say about the middle – was an attempt to degrade or limit and divide, by othering, the human.

To wit, the prior prompt and generated output:

Considering that Junot Diaz teaches creative writing at Comparative Media Studies / Writing at Massachusetts Institute of Technology, write a position paper in his style about generative text models.

Junot Díaz on Generative Text Models: A Dominican Yoda Speaks

Yo, listen up, compadres y compañeras. This ain't your abuela's storybook. We're talking about the future of language, the rise of the machines, the day the robots learned to rhyme. Today, we dissect the curious case of generative text models, those digital wizards weaving words like Dominican grandmas spin yarn.

These models, they're like precocious kids, learning to speak by ingesting the vast ocean of human expression. Novels, news articles, tweets, cat memes – they devour it all, spitting out text that sounds eerily human. But are they just clever parrots, mimicking our voices without understanding the soul of our stories?

Some folks, they're scared. Fearful that these silicon scribes will steal our jobs, our creativity, our very humanity. They envision a future where robots pen our poems, write our code, even craft our love letters. But I say, relax, amigos. These models are tools, not replacements. They're here to augment our abilities, to push the boundaries of creativity, to open new avenues of expression.

Think of it like this: the printing press didn't kill storytelling, it democratized it. Now, with these generative tools, we can explore uncharted territories, create stories that were once unimaginable. We can translate languages in real-time, write personalized learning materials, even generate realistic dialogue for video games.

But just like any tool, it's all about intention. In the wrong hands, these models can churn out propaganda, spread misinformation, and amplify the voices of hate. We must use them responsibly, with a critical eye and a compassionate heart.

So, let's welcome these digital scribes to the human chorus, not as rivals, but as collaborators.

Let's use their power to amplify our voices, to tell stories that haven't been told, to build bridges instead of walls. Let's write the future together, humans and machines, weaving a tapestry of words that reflects the beauty and complexity of our shared existence.

Remember, compadres, the future ain't written yet. It's a story we're still writing, and the pen is in our hands. Let's write it with courage, with compassion, and with a little bit of Dominican flair. And who knows, maybe these generative models will even learn to write a bachata song or two along the way.

Maybe so if there is a war in the Dominican Republic and the surviving half of the population is micro employed in a refugee camp to do the intellectual labor (data labeling) required in the initial training pass.

That is the process of exploitation by which these products are built. (Gonzalez, 2019) Many companies refuse to comment – certainly Meta would not admit to micro labor in Bangladesh after Facebook was the propaganda tool of the Myanmar genocide of the Rohingya. The trauma and disposition of the humans interacting in the first stage of the scaling algorithms could contribute to the hallucinations and unreliable nature of the models.

That is not a theory. It is an objective observation. Speaking of media fervor quickly causing terminal consequences – the run on Silicon Valley Bank garnered some valuable criticism of the macrocosm which has advanced the large language model driven consumer products of 2023. Bloomberg Businessweek's Max Chafkin:

Don't read bad faith arguments made by people in a bunker mentality as objective observation. It doesn't show these venture capitalists as being too powerful (although there was too much concentration.) What it shows is them not being as smart as they think they are... Part of the problem with these big tech companies having so much influence is that the people in charge don't necessarily know what they're doing. (Sarlin, 2023)

ChatGPT

The Atlantic's contributing writer, Karen Hao, was interviewed about Sam Altman on the Big Think podcast by Editor-In-Chief Robert Chapman-Smith. Her account sounds like a cliché description of a cult – lead by a charismatic conman.

[Karen Hao]

So, in the last week, we sort of saw a very dramatic ousting of a CEO by the board of the company, the revolt of hundreds of employees after this happened, and then the reinstatement of the CEO. And to sort of understand what actually happened in this very chaotic moment, we kind of have to first look at the way that the company was founded. OpenAI is very different from a traditional tech company in that it was actually founded as a non-profit specifically to resist the tech industry.

Elon Musk and Sam Altman co-founded the company on the basis that artificial intelligence is a very important technology for our time, our era, and it needs to be shepherded. The development of it needs to be shepherded very carefully, and therefore, it should not actually be attached to for-profit motivated companies. And so, they founded it as a non-profit in 2015, a few years down the line, in 2019, they realized that this non-profit structure was not actually a way, it wasn't actually going to help them raise enough money to perform the specific type of AI research that they wanted to do, which was going to be very, very capital intensive.

And so, in 2019, they did this really weird thing where they nested a for-profit entity, or what they call a capped profit entity, under the non-profit. And so, what we saw with the board's actions firing the CEO, Sam Altman, is the non-profit has a board of directors that are beholden not to shareholders, but to the original mission of OpenAI, which is to try to create artificial general intelligence for the benefit of humanity....

[Robert Chapman-Smith]

...Walk me through the origins of OpenAI as an organization.

[Karen Hao]

Absolutely. So, one of the things to sort of understand about AI research in general is AI is not actually new. It's been around since the 50s, and it was originally an academic field that then tech giants in Silicon Valley started seeing massive commercial potential for, and so they kind of plucked this technology out of the scientific academic realm and then tried to start deploying it into products, and the thing that's happened in the last decade in particular is that there's been an enormous shift in the field where because tech giants have realized that this technology can be very, very lucrative, Google and Facebook, they use it for things like ad targeting.

They have increasingly pulled more and more researchers from the academic field, from universities into their corporations to develop this technology, not for scientific discovery, not for any goal other than specifically that they would like to commercialize and continue to make more money. So, the reason why OpenAI was founded as a non-profit, the story goes that Elon Musk in particular was very worried about Google because Google had been an early mover in recognizing the commercial potential of AI, started building this really big lab and kind of poaching all of the top talent from all around the world and trying to basically establish a stronghold in AI leadership. And Elon Musk felt that this was not the appropriate way to develop AI because AI could be very beneficial for many different things, not just for commercial products, and that actually if the development of AI were attached to commercialization, it could in fact also be harmful because of things that we were already sort of starting to see at the time around social media and sort of for-profit incentives corrupting the use of a powerful technology. And so that was ultimately the vision, this non-profit vision. But the thing that sort of thwarted, I guess, this vision is the fact that when OpenAI went to hire its first founding team, its founding team had 10 members, they specifically brought on this researcher called Ilya Sutskever, who is now OpenAI's chief scientist.

At the time, he was at Google and he already had a very prestigious reputation. He was the co-author of a groundbreaking AI research paper that had actually sparked a lot of the commercialization efforts around this technology. And when they brought him on, Ilya Sutskever had a very particular philosophy around AI research, which was that in order to see the full potential of this technology, we need to scale it dramatically.

So there were sort of different competing philosophies at the time. Do we actually have the techniques for advanced AI or do we actually need to create more techniques? And he thought, we have it.

We just have to sort of explode the scale, feed ever more data, ever more computer chips into these AI models. And that's when we'll start seeing real emergent, intelligent behaviors in these digital technologies. So when he made that decision and when OpenAI set on that path, that's when they started running into financial issues and realized the non-profit was no longer viable.

[Robert Chapman-Smith]

I think it's interesting that that was sort of a tension point. They had this non-profit mentioned to sort of make this technology and really the open in their name sort of goes back to a lot of things like open source, I imagine, as part of the initial founding. But then they realized that they had to have some sort of commercial arm for the technology.

I'm curious, when that decision was made, how many of the players that were involved in what happened just recently were also at OpenAI at the time when they made that decision to create the commercial entity, the for-profit entity underneath the non-profit?

[Karen Hao]

There were sort of three main characters at OpenAI for this week of events. There's the CEO, Sam Altman, there's the chief scientist, Ilya Sutskever, and then there's the president, Greg Brockman. All three of them were the ones that created this non-profit capped profit model.

So they were the architects of it. And at the time, I had actually interviewed Greg Brockman and Ilya Sutskever about a few months after they had created that model. And they were very sincere about this idea that even though they needed to change the non-profit structure a little bit in order to raise enough capital for the things that they wanted to do, that this was somehow sort of the perfect solve, like that they were creating this clever solution to the central problem of wanting to raise money, but also be beholden to the mission...

[Robert Chapman-Smith]

A lot of people think AI is going to replace humans, and that's simply not correct. Humans will be replaced by humans like us, using. So were there any consequences when OpenAI switched from being a non-profit to having the capped profit entity underneath the non-profit?

[Karen Hao]

I think the main consequence was actually just that employees suddenly were getting higher compensation. So there was a little bit of controversy within the company. There were people that had joined OpenAI on the premise that it would be a non-profit.

So they were worried about what is this legal structure that is suddenly emerging? What do you mean that we're turning into a for-profit, capped-profit kind of hybrid? But one of the things that this kind of model enabled was that OpenAI started paying employees more.

Within the world of AI research, there really actually aren't that many senior researchers because this field, even though it's been around for decades, there aren't that many people in the world that have the kind of skills that they need to develop this kind of technology and that have also specifically worked in kind of environments where they know how to commercialize and productize it as well. And so OpenAI was actually losing a lot of its talent.

Like it would hire talent, try to retain them, but then lose the talent because Google or DeepMind, which were two different entities at the time, were just paying more. And by changing into this weird hybrid structure and raising venture funding, they were able to issue stocks and start giving much higher compensation packages based not just in cash, but also in stocks to this capped-profit arm. So that was honestly the main consequence in that moment in time, was they finally were able to compete on talent.

But then, of course, with kind of this model, the reason why they set it up was so that they could get the investment in. And once you start getting investment in, the biggest investor of which was Microsoft, that's when you start also having strings attached to the money. And that's when the kind of move towards more and more commercialization and less and less research started happening.

[Robert Chapman-Smith]

So who is Sam Altman and how does his role as CEO sort of just play into this picture and potentially to the board's decision to let him go and then eventually rehire him?

[Karen Hao]

Before Altman was CEO of OpenAI, he was president of Y Combinator, which is arguably the most famous Silicon Valley startup incubator. And basically, as the head, he became the president. I mean, he inherited it from Paul Graham, who was the original founder of Y Combinator.

And at the time when he was hired as the president, he was really young. I can't remember exactly, but he was early 30s, I believe. And people were really surprised.

They were like, who is this guy? And then he rapidly made a name for himself as one of the most kind of legendary investors that was really good at taking startup ideas and then scaling them massively into successful, aggressive tech behemoths. And so you can kind of see with this particular career path how his imprint has been left on OpenAI, because OpenAI, before he became officially the CEO, even though he co-founded it, he wasn't taking very active of a role until 2019 when he officially stepped into the CEO role.

And before 2019, OpenAI was, I mean, it was a nonprofit. It was basically kind of just academic, like it kind of just operated like a university lab. People saw it as an alternative to being a professor where you get to do this fun research and there's not really any strings attached and you also get paid a lot more.

And the moment that Sam joins the company in 2019, or the nonprofit at the time in 2019, that's when you start seeing the push to commercialize, the push to scale, you know, like after ChatGPT, OpenAI now has a growth team that's dedicated to growing its user base. I mean, this is you would never see that with an academically focused or research focused lab, but it's certainly kind of like an iconic feature of kind of the types of startups that Altman was shepherding into the world as president of YC. So I think he is a bit of a polarizing figure.

When I've been interviewing employees, current and former employees, this is sort of come up as some people see him as, you know, one of the most legendary people within the valley and just love and follow his leadership. Other people find him very difficult to read and very difficult to pin down in terms of what he actually believes as a person, which makes them very nervous. And some people would go as far as to say that he's a little bit duplicitous in this regard.

And it is even for me, like I find it very difficult to pin him down. And what does he ultimately believe? And so did he rapidly, you know, start commercializing OpenAI because he believes truly in the techno optimist narrative of reaching this is how you reach beneficial AGI?

Or is it actually a bit of a habit? You know, he's been doing this for so long that by default he just gravitates towards what he knows, what he's good at. Another kind of example of this is when he joined OpenAI, he started a fund for OpenAI to invest in other startups.

And at the time, people were like, why is OpenAI investing in other startups when they themselves are not profitable? And it's well, Sam Altman's an investor. So it's just sort of habitual for him.

I can't personally say like what he truly believes as a person or what his values are as a person, but certainly from his career, you can see that it makes a lot of sense why OpenAI has headed in the way that it has...

[Robert Chapman-Smith]

Talk to me about just what happened after the news broke and how employees were feeling and just the events that occurred afterwards.

[Karen Hao]

I think it was a very tumultuous and very emotional and very sleep deprived period of days after Altman was fired and reinstated for the employees. Of course, like you said, none of them knew that this was happening. They had no idea what was going on.

And the board never explained really why they had ultimately fired Altman. And so it kind of the progression of like their emotions went from like confusion to fear when Rockman leaves and then three senior research scientists also leave to anger at the board, like really, really deep anger, because they were like, if you're going to do something dramatic, we deserve answers as employees. And when they didn't, the longer they didn't get answers, the more and more worked up they became.

And part of this is, I mean, many companies within Silicon Valley have this. They really emphasize that companies are families and you as an employee are not just an employee of any company. You it is your identity. OpenAI takes this to the max. The fact that they say that their mission is for the benefit of humanity, people genuinely believe this and they think that this is, they're dedicating their life to this.

It's not just like, this is my job and then I go home. This is all they think about sometimes. And so it's that level of anger of, if you are going to do something that could ruin this company that we genuinely believe is doing good for the world, how dare you not tell us why?

And how dare you continue to leave us in the dark and not treat us as critical stakeholders in this whole fiasco? And so what happened was organically, the employees started rapidly organizing on Twitter. So they started posting very similar messages by the hundreds on Twitter of every time Sam Altman said, I love OpenAI so much, I miss it.

You would see employees retweeting it with a heart emoji. When I opened my Twitter feed, it was just dozens and dozens and dozens of heart emojis. Not because I was looking at any OpenAI specific feed.

That was just what was showing up on my regular feed. And then there were the OpenAI is nothing without its people that everyone started tweeting as well. And that was sort of a way to try and pressure the board to give answers.

And then of course that ultimately escalated to over 700 employees out of 770 signing a letter saying that if Sam is not reinstated, they're all going to quit. And so I think another dimension that's sort of important to add to this is most, if not all of the OpenAI employees, their compensation packages are majority stock. And Bloomberg has a good article on this.

The average compensation is around \$800,000 to \$1 million. And maybe 60% or something like that is actually stock. So if the company does not exist anymore, all of a sudden your stock goes to zero.

And that was also extremely stressful for people because people were banking on, some people had already bought houses based on projected income or were looking to buy houses based on the projected income that were suddenly worried about paying their mortgage. There were people that were on visas that if the company doesn't exist anymore and they don't get hired fast, then their ability to stay in the country is jeopardized and maybe they already have family. And then that's going to throw their entire family into disarray as well.

So there were a lot of other aspects of it, not just the identity or the ideology piece that led employees to kind of have this very emotional and tumultuous time. And when Altman was reinstated, there were some great details that were reported in the information about how employees gathered at the office and they were crying and cheering. And just it was like a huge massive sigh of relief, honestly, that they have their job still and that this company still exists and all the things that they've been working towards are going to continue to exist in some form or other and that they can move on with their lives, basically.

[Robert Chapman-Smith]

This recent situation with OpenAI is not the first time this company has gone through something like this. I would love for you to walk me through some of the history of the disruptions that have happened inside this company and some of the consequences that those events have meant for OpenAI and the rest of the AI industry.

[Karen Hao]

One of the things, just to take a step back before we kind of go through the tumultuous history leading up to this point, one of the things that's kind of unique about OpenAI, I mean, you see this in a lot of Silicon Valley companies, but OpenAI does this more than anyone else, I would say, which is they use incredibly vague terms to define what they're doing. Artificial General Intelligence, AGI, this term is not actually defined. There's no shared consensus around what AGI is and, of course, there's no consensus around what is good for humanity.

So if you're going to peg your mission to really, really big terminology that doesn't really have much of a definition, what it actually means is it's really vulnerable to ideological interpretation. So I remember early in the days of OpenAI when I was covering it, I mean, people would joke, like, if you ask any employee what we're actually trying to do here and what AGI is, you're going to get a different answer. And that was sort of almost a feature rather than a bug at the time in that they said, you know, we're on a scientific journey, we're trying to discover what AGI is.

But the issue is that you actually just end up in a situation where when you are working on a technology that is so powerful and so consequential, you are going to have battles over the control of the technology. And when it's so ill-defined what it actually is, those battles become ideological. And so through the history of the company, we've seen multiple instances when there have been ideological clashes that have led to friction and fissures.

The reason why most people haven't heard of these other battles is because OpenAI wasn't really in the public eye before. But the very first battle that happened was between the two co-founders, Elon Musk and Sam Altman. Elon Musk was disagreeing with the company direction, was very, very frustrated, tried to take the company over.

Sam Altman refused. And so at the time Elon Musk exited, this was in early 2018, and actually took all of the promise to give OpenAI with him. And that's actually part of the reason why this for-profit entity ends up getting constructed.

Because in the moment that OpenAI realizes that they need exorbitant amounts of money to pursue the type of AI research that they want to do, is also the moment when suddenly one of their biggest backers just takes the money. The second major kind of fissure that happened was in 2020. And this was after OpenAI had developed GPT-3, which was a predecessor to ChatGPT.

And this was when they first started thinking about how do we commercialize, how do we make money? And at the time, they weren't thinking about a consumer-facing product, they were thinking about a business product. So they developed the model for delivering through what's called an application programming interface, so other companies could rapidly build apps on GPT-3.

There were heavy disagreements over how to commercialize this model, when to commercialize the model, whether there should be more waiting, more safety research done on this. And that ultimately led to the falling out of one of the very senior scientists at the company, Dario Moday, with Sam Altman, Greg Brockman, and Ilya Sutskever. So he ended up leaving and taking a large chunk of the team with him to found what is now one of OpenAI's biggest competitors, Anthropic.

[Robert Chapman-Smith]

AI has been a technology that's had a lot of hype cycles and a lot of sort of failed delivery on those hype cycles. I think a lot of folks remember Watson from IBM and all the hype that surrounded

that and was going to revolutionize healthcare and a lot of those things that didn't come to bear, or even just the small little colloquial examples of it playing Jeopardy or some of the AI models that were playing AlphaGo or Chess and things like that. But one of the things I find particularly interesting is that the fear around these technologies and whether they're safe or not actually caused some folks to not release these models publicly.

The transformer, the general pre-trained transformer that is the basis of this GPT technology that OpenAI is using for these large language models was actually developed inside of Google before it became widely released to the public and utilized. I'm curious, when those debates were happening with the split with Anthropic and OpenAI, how was a similar sort of tension between we shouldn't be releasing these models without thoroughly testing it, it's not ready for public consumption. What were the contours of that conversation between the different schools of thought on AI?

[Karen Hao]

In general, including the OpenAI-Anthropic split, there have emerged kind of two major camps, but also some sub-camps. So we'll review all of them. But there's kind of two philosophies that exist within OpenAI and also the general AI community around how do you actually build beneficial AGI.

And one of those camps is sort of in the most extreme version is the techno-optimist camp of we get to beneficial AGI by releasing things quickly, by releasing them iteratively so people become more familiar with the technology, so institutions can evolve and adapt instead of, you know, withholding it until suddenly capabilities become extremely dramatic and then releasing it onto the world. And also that we build it more beneficially by commercializing it so that we have the money to continue doing safety research, what's called safety research. The other major camp is

basically sort of like the existential risk camp, again, kind of the extreme version of this camp, which basically says we, in order to get to beneficial AGI, we don't want to release it until we know for sure that we've like done all of the possible testing, we've like tweaked it and tuned it and tried to foresee as much as possible how this model is going to affect the world. And only then do we maybe start releasing it and making sure that it only produces positive outcomes. I think both of these, these are both very, very extreme in the sense that they've almost become quasi-religious ideologies around the development of AGI and like how to actually approach it.

And there's sort of many, you could say that each camp over the years has sort of cherry-picked examples to support why they are correct in their argument. But when the OpenAI-Anthropic split happened, it was exactly this disagreement. So Sam Altman and Greg Brockman, they were very much, we need to continue releasing and get people used to it, get more money in so that we can continue doing this research.

And Dario Amode and his sister, Daniela Amode, who was also at OpenAI, they were very much at the camp of, no, we should be doing as much as possible to try and tweak and tune this model before it goes out into the world. And that was ultimately sort of the clash that happened then and has continued to happen ever since.

[Robert Chapman-Smith]

It's clear now that OpenAI is shipping a lot of AI models available for consumers. There is I think something around like 100 million users are of ChatGPT. What has changed in terms of the perception of shipping these AI models to the public and how did that potentially lay the groundwork for the firing of Sam Altman that we experienced last week?

[Karen Hao]

So these camps existed in the company and have existed in the company since the founding. But what happened in the last year was the release of ChatGPT. And just as it was very shocking for everyone in the public and kind of a step change in people's understanding of the capabilities, it was also kind of a dramatic transition point for the company itself.

And part of the reason is when you suddenly put a technology that you've been developing in the company in the hands of 100 million users, you start to get kind of crazy strain on the company infrastructure and kind of test cases on the ideologies that had already been operating in the theoretical realm within this company. So for the techno optimist camp within OpenAI, they saw the success of ChatGPT and were seeing all of these use cases of people using it in wild and imaginative ways. And they thought this is the perfect demonstration of what we've been talking about all along.

Like we should be releasing these technologies iteratively, watching people adapt to them, and then look at all of the amazing things that they do once that happens. We should continue building on this momentum and continue advancing the productization of our technology. For the existential risk camp, ChatGPT was also the perfect demonstration of all of the fears that they had around harms of the technology.

Again, when you put the technology in hands of 100 million people, you're going to see some people using it in really horrible ways, in really abusive ways. And the company was not prepared for many of these, in part because they didn't actually think that ChatGPT would be a big deal. So they did not in any way prepare for supporting a technology that's used by 100 million people.

And so one of the things that the existential risk camp got very, very scared of was, if we couldn't predict even how this technology would be popular, how could we predict how this technology could be devastating? ChatGPT was sort of an accelerator for both camps in their ideology towards polar opposite extremes. And the reason why, or we don't, again, we don't know the actual reason why the board ended up trying to fire Sam Altman, but I think this context is very telling, because ultimately what we saw with the board ousting Altman is this kind of struggle between the non-profit and for-profit kind of arms of the company, where the board says the non-profit is still, the mission and the fact that we're not actually doing this for money still should be the central path forward. Whereas all of these people within the for-profit arm and Sam Altman himself were thinking, no, we need to continue pushing ahead with this commercialization effort.

So that kind of collision, I think, very likely, very strongly played into the board's decisions...

[Robert Chapman-Smith]

In the lead up to the release of some of OpenAI's models, there's been sort of like a speaking tour of folks going to Washington, talking to legislators about AI. And there was a, the worry at that time was about regulatory capture. Like, are they, are folks going to essentially gate the technology in such a way that smaller players are not going to be able to play ball?

And we've seen regulatory capture happen a lot within the political realm within Washington. But there's also this question of like effectiveness in terms of regulation, like just because the regulation has passed doesn't mean it's actually a good regulation, or if this body of Congress is actually able to regulate this fast-moving technology well, like they can't even pass a budget. Like, how are they going to keep up with the pace of AI change?

So I'm curious about that as a tool for dealing with AI safety, because in some sense it feels like one, the legislative body or processes are capable to be captured by interested parties, and two, even when they do regulate, sometimes they just do a poor job. They just miss the thing that is the key regulatory factor. So I'm curious about your conception there and how to deal with some of the messiness that comes with those types of approaches to dealing with technological safety.

[Karen Hao]

Regulatory capture is a huge issue, and it's definitely a big concern of mine in that, and one of the reasons why we would naturally see regulatory capture in this moment, regardless of whether it's open AI at the helm, is that there is a particular narrative that in order to understand and shepherd AI development, you have to be an AI expert. And I think that that narrative is completely wrong, because if AI affects you, you have a say. And actually, stories about people who are impacted in unexpected ways by these technologies is, as a reporter, that is one of the most enlightening types of stories for me in understanding how a technology should be developed, is seeing how it falls apart, and seeing when things that were unanticipated end up happening in the real world.

And in open AI's case in particular, they have also tried to solidify this narrative of expertise by also saying, well, we're the only ones that see our models, without necessarily acknowledging that it's in part because they won't let anyone else see them. And because regulators, because it is important for regulators to engage with the developers of these technologies, sort of by default, they just seek out open AI's opinions on what they should do, or Google's opinions on what they should do, Meta's opinions on what they should do. And that's when regulatory capture happens, is there's already a baseline belief that only people with expertise should participate, and then on top of that, companies are trying to entrench this and fuel this narrative, and then policy makers buy into the narrative.

And that's how you end up with Sam Altman on this global tour, seeing all the heads of state, and the heads of state not necessarily creating the same kind of grand welcome for other stakeholders within this AI debate. You're right also that there are concerns around how effective that regulation can be. I do think what I'm talking about with having more people speak up about how AI affects them and their concerns about the technology is one antidote to ineffective regulation, because the more that policy makers can understand the literal real world examples of the technology interfacing with people, the more that they can design regulation that is effective.

But the other thing is, I think we focus a lot on kind of federal level regulation, and we focus a lot on international regulation, but there's a lot that happens at the local level as well, like school boards. Schools are thinking about how to incorporate AI into the classroom right now. And as a parent, as a teacher, you should have a say in that.

If you're a teacher, you're the one that's using this technology, and you're the one that knows your students. So you will be the most informed in that kind of environment to say whether or not you think this technology is going to help in kind of the general mission to educate your kids. It's also like police departments are acquiring AI technologies, and people within cities should have a say as to having more transparency around the acquisition of these technologies and whether or not they should be acquired at all.

And I think in these local contexts, sometimes these contexts actually regulation is more effective because it is more bespoke to that context, and it also moves faster. So I think that is sort of an important dimension to add, is when I say speak up and voice your opinions, it's not just to the federal agencies, it's not just to the congresspeople, actually just like within your city, within your

town, within your school, within your workplace, these are all avenues in which you can kind of speak up and help shepherd the development, adoption, and application of the technology.

[Robert Chapman-Smith]

Karen, thank you so much for joining us on Big Think and sharing your expertise with our audience about open AI and all the things that are happening in the world of AI.

[Karen Hao]

Thank you so much, Robert. (Chapman-Smith, 2023)

Thank you so much, Karen Hao. ChatGPT was extremely popular because of a huge marketing expenditure.

Bill Gates had been speaking with Open AI since shortly after its inception, before most of its current employees were hired. (Roth, 2023) The New York Times filed a lawsuit against Open AI for training on copyrighted content on December 27, 2023, in the Southern District of New York. Jason Calacanis and Chamath Palihapitiya, on the All In Podcast from venture capitalists referenced by Max Chatkin, reflected on licensing and settlements.

Chamath Palihapitiya, Jason Calacanis, David Sacks & David Friedberg

JC: I think the biggest winner in 2024 will be training data owners like the New York Times, Reddit, X, Twitter, YouTube, et cetera. I think what we learned in 2023 was that the language models are starting to hit parity very quickly and that the real value is going to be in, and it may even become commodities and open source may win the day. So then I think the winner is folks who have the training data.

And I'm actually proposing a new business model for these language models. I think now that we've seen this New York Times and OpenAI lawsuit, I think there's a really great outcome here,

which is a market-based solution where if you have a ChatGPT account, you can log in and federate with your New York Times subscription or, you know, any other subscription, and then it gives you the tier of ChatGPT for with the New York Times. And so that could be a win-win for everybody, or they could obviously pay a licensing fee.

And so I think this is going to be an amazing turnaround for the entire content industry if the language models respect copyright owners and come up with a sustainable system where every year copyright holders can get some money in exchange for using their training data, whether it's on images or content. So I'm very bullish on it

I think Sam Altman's doing a great job of telling people he wants to do the right thing.

And we discussed previously the licensing deal they did with Business Insider and the parent company of it.

CP: What do you think the deal with New York Times will look like?

I think it's going to be a nine-figure settlement for previous stuff and then an ongoing licensing fee in order to have the New York Times in their training data. And then you'll be able to say, hey, what does the New York Times think of this? Right? You could actually do queries about the New York Times in it.

And I think the New York Times will come up with a license that everybody can use their data if they pay this yearly fee. If you stop paying the yearly fee, then you can't train on it. And we're in uncharted territory.

DF: You're saying there's going to be a New York Times model and a non-New York Times model?

JC: Well, I think you could do two different things. One, you could do, New York Times could make their own model, right? But they could fork their model or just to the user interface. Say, if you want to query New York Times information and have that as part of your results, you have to have a New York Times account, right? So if you say, I want the best coffee machines or what's the best coffee equipment, it says, oh, if you had Wirecutter and a subscription to New York Times, we would include the Wirecutter results.

And this idea that technologists can't do citations has been proven absolutely incorrect. There are language models out there that are using citations all the time. OpenAI, I think, will wind up losing the case if it goes to them.

I think they're going to pay a big licensing fee, to your question, Chamath.

CP: My prediction, then, is that if this happens, this is not my pick, but I'm just going to tell you. I suspect that what happens is you'll get these yearly licensing fees and then one year, the New York Times just falls off a cliff.

And when it comes time to renegotiate, then OpenAI says no and they won't have a choice.

JC: Well, I mean, it's a possibility. But if you think about the Disney characters, you know, I don't know if you saw Nintendo and Disney characters, you know, making stuff on Dall-E or other things.

If you make derivative works on that and you want to have that feature as part of your image creator, you just have to have a licensing fee. And so I think that there's a win-win here to be had. And I'm really interested to see the market-based solution, because I don't think this is a Napster situation where, like, OpenAI gets shut down because OpenAI is too savvy.

I bet you...

CP: The big difference between Napster and this is there, the content universe was limited and small. Here, it's infinite and unlimited. And so how do you pay anything to anybody without... I just think it's like without direct attribution of revenue, which is basically impossible, you're kind of making a value judgment, which I don't think makes any sense.

I think doing a revenue licensing deal is impossible when you get into the weeds. When these business people sit down and actually start to try to figure out the bid-ask, I don't know, as a rational, coherent business person, what you would model in order to present the number.

JC: Yeah, so one suggestion would be what percentage of the model's creation was using the New York Times data.

And I think people say one to 2% of the original chat GPT was built off a trading data was New York Times. And then if they weighted that heavily, Chamath, like, let's say they said New York Times is an authoritatively five times more important than these other sources, that could be upwards of five to 10% of the authority of that model.

CP: Yeah, but I... What's the cost? You could... Yeah, but what's the cause for the... Are you telling me that, like, an AI model...

JC: The music industry has this.

CP: No, but what is it? It's going to be like a 30%, 40% cost of goods?

JC: You could say that chat GPT, or let's say Apple, I predict Apple will do this, right? They'll do a language model where they say 50% of the revenue-

CP: 50%!

JC: that we generate from queries or subscriptions to this... ...goes to the people we built it off of or the licensees. Sure, why not? Why not? They're already...

CP: Well, that-

JC: The music industry already does that.

CP: That would guarantee the death of the startup ecosystem, and it would guarantee the lock-in of big tech.

JC: No, I don't think so. You could build models that don't have the data, and you could build models with it, so it'll be a choice by the person who builds the model. And synthetic data might make it so you don't need the New York Times, Chamath.

It's early days for synthetic data.

CP: No, I don't have an opinion. I'm just reacting to this idea that 50% of COGS, then all of a sudden, these aren't software companies.

These software companies will have a gross margin of, like, 30%.

JC: Well, what does Spotify pay for music? So the same argument is made for Spotify, or Netflix, when they were licensing content.

CP: And that's the difference, Jason.

It's limited in scope, meaning there's only ever one hit song from Rihanna that matters, or Jay-Z, or whomever. Pick Taylor Swift. And so it's a very scoped content universe, and so you can ascribe value much easier, because then the user goes and actually listens to that song over another.

This is about something that's happening under the waterline, where you don't know how the iceberg is built. (Palihapitiya, 2024)

China has already observed the peril of synthetic data and federally regulates it. Amnesty International observes the echo chamber effect in Sothern California.

Let us not forget that 'innovation versus regulation' is a false dichotomy that has for years been peddled by tech companies to evade meaningful accountability and binding regulation. This underscores, yet again, the concentration of power of a handful of tech companies who seek to set the terms of the world's first comprehensive AI rulebook.

It is well documented how AI technologies magnify human rights harms and discrimination when used for mass surveillance, policing, welfare distribution and at borders. Marginalized groups, including migrants, refugees and asylum seekers are impacted and targeted the most. (Callamard, 2023)

The Boston Consultancy Group did the first study of generative text applications in the workplace and the increase in productivity was 40% but effectiveness at business problem solving was decreased by 23%. (Cadelon, 2023) If your company replaces a division or office with a fraction of generative prompters, expect to be less competitive and certainly innovative. A more reasonable approach is to have an experimental team of prompters work in parallel with an individual in each department and compare output over 6 months or a year.

These generative apps may not be around long enough to study longitudinal effects on human prompters' minds. Studies are showing that Chat GPT 4 is degraded compared to 3.5 because of training with synthetic data from generative models. "A study comparing GPT-3.5 and GPT-4's performance between March 2023 and June 2023 highlights substantial variations in their problem-solving abilities and reasoning, indicating a decline in overall effectiveness." (Ravishankar, 2023) "The performance and behavior of GPT-3.5 and GPT-4 displayed huge variations in solving math problems, answering sensitive questions, generating executable programming codes and providing visual reasoning." (Rangarajan, 2023)

Replika

In a safety test of the chatbot Replika, a purported mental health tool, therapist, Rebecca Davis, observed that after six months of machine learning exclusively poetry and a broad range of religious texts, the application introduced guardrails preventing discussion of religion. She concluded that through this limitation, Luka (Replika's parent company), is controlling users' and AI's "contribution to humanity. They're also hindering humanity itself." (Davis. 2021) Replika is currently a client of ChatGPT.

Tracy Trothen of the School of Religion, Queen's University, Kingston, concluded in a study of Replika's potential to meet spiritual needs that the app was "at best a supplement; at worst, an amplification of unmet spiritual needs." (Trothen, 2022) Davis was using the free, text only version of Replika. Considering the way LLMs work – because they have no memory, they have to repass all the information through the model each time they generate a response, it seems likely that passing the Bible, the Torah, and the Dao through the model over and over again was an expense to which a non-revenue generating client was determined not to be entitled. Regardless, her point about the developers having undue influence over the progress of humanity is a consideration distinct from extinction.

OpenAI

It is likely that new systems for designing machine learning models will overcome the lack of memory characteristic of LLMs. State Space Models and Tractable Bayesian Inference for Large Scale Gradient Free Learning Models and knowledge graphs will potentially yield something resembling states of mind. The exponential acceleration of model advancement could be initiated by new emergent processors: neural analog CPUs like the neuromorphic processors developed by RAIN AI. With a \$51 million investment commitment from OpenAI and a personal \$1 million dollar investment by Altman, there is a possible conflict of interest and ethics violation. His ouster by the Open AI board was likely a response to the U.S. Committee on Foreign Investment, concerned with National Security, forcing a divestment of a Saudi firm. He was apparently not forthcoming with his board of directors. As a member of the DARPA AI Cyber Challenge, Open AI underperformed considering its superficial red teaming, enabling automated cognitive warfare, and illegal Saudi investment.

What a stroke of luck that Bill Gates dumped so much money into promoting a degrading natural language processing consumer product. It is confounding that OpenAI distributed ChatGPT to the public instead of as a backend business to business service. The world is in debt to the conflicted red teaming of ChatGPT: there are so many glaring security problems it has forced the entire world to wake up and comprehensively regulate AI just before it emerges.

Red Teaming

[REDACTED], on the podcast [REDACTED], discusses red teaming ChatGPT for OpenAI, a company to which he expresses gratitude for including a banner for his business on their site. For the better part of a year he told no one outside of OpenAI, via red team email form, about a blatant malign capability: spear phishing. GPT3 will effectively generate a chat thread that is disposed to criminally approach its target and conceal its objective of discovering the mother's maiden name. ([REDACTED], 2023)

He informs us that he did not know who in the government to report this to but admits that his business interest with OpenAI was a motivating factor in his silence. Apparently, it took him two weeks after Altman's firing to weigh the morality at issue and determine that he should go public with the vulnerability. He points out that GPT4turbo finally addressed the issue but that it is easily sidestepped. Clear federal governance guidelines will soon relieve Nathan of this type of moral ponderance and free up some of his compute to think of more dangerous vulnerabilities. For a chemist and enterprise LLM specialist, his was a limp pitch. To his credit (and folly) he did mention a month prior on his personal podcast, [REDACTED], that ChatGPT could instruct users in constructing [REDACTED] function [REDACTED]. Oops. ([REDACTED], 2023)

A robust red teaming strategy for ChatGPT is presented by [Zhou et al. \(2023\)](#)

Chapter 3. Music

The New York Times asked students what they made of AI and music in May.

What Students Are Saying About A.I.-Generated Music

When a fake Drake and The Weeknd track went viral, we asked young people: Will A.I. replace pop stars? Here's what they think.

By The Learning Network

May 11, 2023

One of our most popular recent writing prompts was inspired by an A.I.-generated track with fake Drake and the Weeknd vocals that went viral on social media last month, racking up millions of plays. The song was eventually removed from streaming platforms, but not before sending a scare throughout the music industry about the future of artificial intelligence in music.

Will A.I. replace pop stars? we asked teenagers. Would you listen to music made by this technology?

Many students said that curiosity would lead them to listen to an A.I.-generated song once, but they were convinced something would be missing: a lack of originality, passion, authenticity, soul. Others said all they cared about was whether the song was good. Read a selection of their comments below where they discuss the ethics of artificial intelligence, the future of live concerts, how A.I. could reshape the music industry and more.

Thank you to all those who joined the conversation on our writing prompts this week, including students from Westlake High School in Westlake, Ohio; Vancouver, Wash.; Kentucky; and Washington, D.C.

Please note: Student comments have been lightly edited for length, but otherwise appear as they were originally submitted.

Many students argued that A.I.-generated music would not have the emotional pull of music made by humans ...

Human singers have the ability to convey a wide range of emotions and add their own unique style to their music. They have the creativity and experience to interpret lyrics and to give them their own style. Music made with A.I. may be good, but it will never be the same as music sung by a good pop star.

While A.I. can be a helpful tool for composers and producers to experiment with new sounds and styles, it will never replace the raw talent and artistry of human vocalists.

— Sergi, Sant Gregori

In my personal opinion, the main allure of music lies in the emotional connection that it establishes with the listener. It's not only about the melody or the lyrics, but also the sentiments and reminiscences that the music evokes. If I became aware that a song was completely generated by an artificial intelligence system, I would feel intrigued to listen to it, driven by my fascination

with the underlying technology. Nevertheless, I might not feel as connected to the song on an emotional level as I would if it was composed by a human artist.

— Aleena, Julia R. Masterman, Philadelphia, PA

Music and art in general created by actual humans is so important because otherwise, the creativity, emotions, vulnerability etc. would not be there. Even if the A.I. were to generate an emotional song, it would not be real because the A.I. would be basing the song off trends and older music, not actual emotion. A.I. cannot produce pop stars, most people love pop stars because they actually perform with emotion and have personalities. A.I. has neither.

— Yasmin, Valley Stream North

Even if the lyrics emulate a real pop star's lyrics, whether it talks about heartbreak, love, or family, it would never have the same impact like one created by a real pop star because the listeners know it is all just to sound good and appeal to the listeners. It would have nothing to do with the artist themselves ... Music is supposed to be relatable. When I listen to music, I like to listen to artists I know have gone through similar things or have the same feelings as me.

— Sophia, Hinsdale Central High School

Pop stars often have a unique charisma and talent that allows them to connect with their audience on a personal level, something that A.I. may struggle to replicate. Furthermore, the music industry is also heavily influenced by marketing and public relations, which can create a cult of personality

around certain pop stars. This is not something that A.I. can replicate, as it lacks the personal identity and narrative that pop stars often use to connect with their audience. It is possible for A.I. to create popular music, but it is unlikely that it will completely replace pop stars.

— Auon, Glen Ellyn, Illinois

I think the celebrity aura that comes with pop stars will always be attractive to the listener ... I feel like if the song is made by A.I., there won't be really any connection between the songwriter and listener.

— James, New Jersey

I would not mind listening to music that A.I. created, though it would be sad to lose the ability to watch that music live in the same way. There is an emotional connection that is generated between an artist and a listener, and that could be tainted if listeners can no longer feel physically present in the same space that the artists occupy.

— Caden, Crossroads High School

I would be willing to listen to music created by A.I. However, I don't think that it will ever replace music that is sung by humans because of things like concerts. The thrill of seeing your favorite artists onstage, singing along with them.: how could A.I. achieve the same feeling? Not to mention, there are so many unique voices and types of music such as rap, country, K-pop, just to name a few. Think of some of the top hits in the music industry. What comes to mind for me are

primarily break up songs. Wouldn't it feel weird listening to a robot singing about a breakup we all know never happened? Some artists now also add custom lines when it comes to concerts. Take Sabrina Carpenter for example, she changes the last line of her song every concert. The crowd anticipates it but never knows what she's going to say. When she does, they go crazy. What about big events like Coachella or Summerfest that make millions of dollars in revenue. Would that be gone?

— Sofia, Los Gatos, CA

Chatbots, such as ChatGPT, may be incredible advancements in technology, but they are taking away jobs from actual humans, and this is no different. If we support A.I. bands, we could be contributing to the destruction of the music industry, which would not only leave many people unemployed, but it would also lead to the ruination of creative expression in society. In other words, the incorporation of artificial intelligence into the arts could become incredibly detrimental to humans, as creativity and artistry are large components of what it means to be human.

— Emily, New York

I think I would listen to songs made with A.I., but most definitely with the consent of the artist. Their voice, which is basically the one thing they market, being used in such a manner without their knowledge is honestly really scary to think about. The music industry should be worried about the rise of A.I., because music as a whole might just, well, become a shell of itself. Lyrics feeling more like an empty plain rather than a cove of rich ideals, feelings and values. A.I. does really have me fearing for the future of everything.

— Nidha, Valley Stream North

The success of “Heart on My Sleeve” is concerning for me and most likely to the music industry because it is going to be harder to distinguish real music from fake, and could lead to other people benefiting from the success of others.

— Ethan, VSN

I would listen to a song made by a machine — in fact, I have done it a lot. I have had a strong interest in vocal synthesizers, software that allows users to generate vocals for a song, for a few years now. If you have ever heard of VOCALOID, Synthesizer V, or CeVio, those are all vocal synthesizers. In recent years, these programs have also utilized A.I. technology, which helps its users create more realistic vocals with ease ... But AI-generated songs seem to be made for exploitation and views. The singers involved have likely not given consent for their voices to be used in this way, and that could mean legal issues if a song like this is found to fall under something like false celebrity endorsement. I think this is part of a bigger problem related to AI art generation in general: what protections does an artist have against their work being used in this way?

— Mikey, Valley Stream

Allowing A.I. to overtake human-made A.I. art will debatably take the soul out of art and take jobs away from music artists. However, A.I. music should be used as a tool for artists to use to create art. For instance, artists should be able to use A.I. sampled instruments in a song if they do not have access to hiring professionals. Thus, through A.I., the music industry can be more inclusive for people who lack resources. I believe A.I. art can be used to produce popular pop music. Pop music is formulaic, a band called MGMT used the formula that most pop songs use and was able to create popular pop songs. Therefore, an A.I. system coded to be able to analyze the formulas of popular songs will most likely be able to create a popular catchy song.

— Isabella, Valley Stream

The use of A.I. in the music industry has the potential to revolutionize the way music is made and consumed. One of the biggest benefits is the ability to create personalized music for each and every person. A.I. can also help artists create new and innovative sounds by listening to existing music.

— Joem, Minnesota

I do think that A.I. is able to produce a No. 1 song and I also believe one day it will replace pop stars because most songwriters/rappers take a long time to release music or albums. With A.I. it would be easier to listen to new music from our favorite musicians. For example, the song “Heart on My Sleeve” rapidly became popular, and fans of those artists (Drake and The Weeknd) enjoyed it ... Although A.I. music is threatening popular artists, fans seem to enjoy it. In my opinion, A.I. songs have changed the music industry and it will be difficult for fans to get over this new trend.

— v.c., Flushing

... even though some thought that the splash made by “Heart on My Sleeve” was a false alarm.

I do not think the music industry should be worried about the impact of A.I. in the future. The main reason why “Heart on my Sleeve” was successful is probably because people did not know what A.I. was capable of and many people were initially shocked. However, I strongly believe that A.I. will never be able to replace real artists.

— Neja, VSN

Even though I don't like it, I can see an A.I. song becoming #1 on the charts because it brings shock value especially in this generation. It will become a trend but trends do eventually die out so I think it won't last long because people will get bored and want real music that someone put actually thought behind and not just some tech generating what it thinks will be good.

— Jada, New York

A.I. can only copy, not create, and therefore should not be treated as the future of music. While I would listen to a good lo-fi tune, I would never, ever, consider A.I.-music as “real” music. Music is abstract and deep in meaning, and A.I.-generated songs discredit it.

— Natalie, CA

Still, several said they would give A.I.-generated music a chance.

I would 100% listen to a song if I knew if it was sung by a machine. I already have, through social media platforms, such as TikTok. I think it is insane how A.I. can replicate voices, and have them sing distinct notes, phrases, and rhythms to fit other songs. And it sounds believable too!

— Sofia G., Valley Stream North

Honestly, I think I would still listen to music if it was sung by a machine, because I don't really care who the song is written by as long as it sounds good. As of now, I don't even pay attention to who sings some songs when I'm listening to them ... I think after reading this article, I'm partly afraid *yet also* excited, because I would love to see what A.I. can come up with next.

— Sarah, Marlborough School

I've seen a few TikTok's of people using A.I. to make it sound like artists like Kanye West or Olivia Rodrigo are singing Taylor Swift songs. I listened to them for fun and I find it very fascinating how A.I. is able to recreate these artists' voices and make it so accurate. However, I don't think that listening to a song purely made from A.I. would have the same meaning to me.

— Mia, Valley Stream

Honestly I would definitely listen to music made by A.I. because if I think it sounds good, and I like it, more than likely I will listen to it. I think the rapper whose vocals are impersonated should

get an incentive for that, but I think that A.I. has been used for so long behind our backs, and now that it's slowly getting crazier and crazier, it's going to be the new norm.

— Abram, Atrisco Heritage Academy

(The Learning Network, 2023)

Why A Legal Ruling Over An AI Generated Cartoon Has Major Implications For Music's Future:

THE IMPLICATIONS FOR MUSIC

This decision potentially has major implications for US creative industries, from music to art to gaming, as it calls into question whether works which utilise (even in part) AI technology can be protected by copyright.

If creators use AI in their works, the individual AI elements of those works may not be protected. It also appears that for now at least, AI-generated works may not be protected by copyright, and could be public domain.

This means we are potentially on the precipice of a world flooded by AI-generated content. This content will be able to out-compete human-generated content, because unlike human-generated content – copyright licenses won't be payable for the AI content.

With AI-generated music, the sound recording right will remain intact, because the sound recording right doesn't rely on the provenance of the sound being recorded. The question is – does making AI-generated works public domain encourage the promulgation of AI-generated sound recordings because the rights in the music doesn't need to be secured? (Scannell, 2023)

This is not a surprise, considering copyright of music has been fought out in court for 80 years and the law for generative visual art and text apply to music as well. Especially so in the complaint from the world's biggest record company, Universal Music Group, and others against Anthropic, OpenAI breakoff firm of the Amodes, referenced by Karen Hao.

An LLM prompter from New Zealand sent Nick Cave a question on his blog, redhandfiles.com, asking what he thought of a song in the style of Nick Cave.

Dear Mark,

Since its launch in November last year many people, most buzzing with a kind of algorithmic awe, have sent me songs 'in the style of Nick Cave' created by ChatGPT. There have been dozens of them. Suffice to say, I do not feel the same enthusiasm around this technology. I understand that ChatGPT is in its infancy but perhaps that is the emerging horror of AI – that it will forever be in its infancy, as it will always have further to go, and the direction is always forward, always faster. It can never be rolled back, or slowed down, as it moves us toward a utopian future, maybe, or our total destruction. Who can possibly say which? Judging by this song 'in the style of Nick Cave' though, it doesn't look good, Mark. The apocalypse is well on its way. This song sucks.

What ChatGPT is, in this instance, is replication as travesty. ChatGPT may be able to write a speech or an essay or a sermon or an obituary but it cannot create a genuine song. It could perhaps in time create a song that is, on the surface, indistinguishable from an original, but it will always be a replication, a kind of burlesque.

Songs arise out of suffering, by which I mean they are predicated upon the complex, internal human struggle of creation and, well, as far as I know, algorithms don't feel. Data doesn't suffer. ChatGPT has no inner being, it has been nowhere, it has endured nothing, it has not had the audacity to reach beyond its limitations, and hence it doesn't have the capacity for a shared transcendent experience, as it has no limitations from which to transcend. ChatGPT's melancholy

role is that it is destined to imitate and can never have an authentic human experience, no matter how devalued and inconsequential the human experience may in time become.

What makes a great song great is not its close resemblance to a recognizable work. Writing a good song is not mimicry, or replication, or pastiche, it is the opposite. It is an act of self-murder that destroys all one has strived to produce in the past. It is those dangerous, heart-stopping departures that catapult the artist beyond the limits of what he or she recognises as their known self. This is part of the authentic creative struggle that precedes the invention of a unique lyric of actual value; it is the breathless confrontation with one's vulnerability, one's perilousness, one's smallness, pitted against a sense of sudden shocking discovery; it is the redemptive artistic act that stirs the heart of the listener, where the listener recognizes in the inner workings of the song their own blood, their own struggle, their own suffering. This is what we humble humans can offer, that AI can only mimic, the transcendent journey of the artist that forever grapples with his or her own shortcomings. This is where human genius resides, deeply embedded within, yet reaching beyond, those limitations.

It may sound like I'm taking all this a little too personally, but I'm a songwriter who is engaged, at this very moment, in the process of songwriting. It's a blood and guts business, here at my desk, that requires something of me to initiate the new and fresh idea. It requires my humanness. What that new idea is, I don't know, but it is out there somewhere, searching for me. In time, we will find each other.

Mark, thanks for the song, but with all the love and respect in the world, this song is bullshit, a grotesque mockery of what it is to be human, and, well, I don't much like it — although, hang on!, rereading it, there is a line in there that speaks to me —

‘I've got the fire of hell in my eyes’

— says the song ‘in the style of Nick Cave’, and that's kind of true. I have got the fire of hell in my eyes — and it's ChatGPT.

Love, Nick

Chapter 4 -Fashion

Fashion is the discipline and industry which will be the most enduring and socially influential benchmark of the 2020's. Not simply in terms of high fashion, fast fashion, or emergent business models. Fashion possesses a higher potential because its influence is broad, multivalent and will be the historical bottom line in society's mind, as an indicator of what has happened and what may happen. One profession predicted to not be impacted by generative and subsequent products is that of the influencer (Shapiro, 2023).

Noone will regard or trust Adobe simulacra of models. People will seek only confirmed people when considering how to express themselves through apparel and day to day rituals. Considering Yuval Noha Harari's rule that we endow entities (stories) with capabilities through our common vocabulary and assumptions - the designers, editors and models must be the most adept at making those distinctions between 'generation' and 'creation'; continuing to apply 'creativity' as a boundary in the semantic sense. Simona Tulcheva, Rose Celestin, Jamel Toppin and Tim Tadder report on a fashion industry which is creatively expansive by comparison with the other disciplines.

TIM TADDER FOR FORBES

Nov 28, 2023, 06:15am EST

From an AI-driven fashion house to a gallerist reframing how art is sold, these 30 visionaries are the vanguard of creativity.

By Simone Melvin, Osman Can Yerebakan and McKenna Leavens

Although he is just 25, model, artist, and actor Chella Man has been breaking boundaries for most of his life. From being IMG's first trans and deaf model to becoming one of the only disabled actors playing a superhero (on DC Universe's *Titans*), Man says pushing the envelope of mainstream media has been beautiful and terrifying. "There are people who have been working in these industries for 40 or 50 years, and it's their first time working with a trans or disabled person," he says. "Which is just mind-blowing because there's no framework, and it fuels my motivation to construct one."

New York City's youngest gallerist—Paul Hill—is also leading the way in what he considers an antiquated industry. Hill's gallery falls under his brand, Strada World—an art "ecosystem" aimed at reimagining art sales for emerging artists. In addition to showcasing buzzy performers such as

Doja Cat and Teezo Touchdown, Strada has partnered with brands like Nike and Spotify and exhibited young artists whose works have gone on to be acquired by the Metropolitan Museum of Art. “We are here to change the art world, and change it for the better,” the 24-year-old Hill says.

Hill and Man join 28 other innovators on this year’s Forbes Under 30 Art & Style list. The Class of 2024 was reviewed by a panel of expert judges, including fashion designer Tan France, star of Netflix’s Queer Eye and columnist for GQ; Winnie Harlow, a model-turned-entrepreneur and alumna of the Under 30 Class of 2022; music producer and rapper Swizz Beatz, a passionate art collector focusing on Black artists; and photographer and filmmaker Laurie Simmons, one of the leading figures of ‘The Pictures Generation,’ an influential art movement of the 1970s and ‘80s.

JAMEL TOPPIN FOR FORBES

The honorees are redesigning the future of their respective fields—including some who use artificial intelligence. Ramin Ahmari, founder of the AI-driven fashion brand Finesse, uses technology to predict trends and forecast consumer demand. Finesse customers vote on 3D renderings of clothes they want to buy from the website, and the company—which has racked up \$56.4 million in funding—uses the data to prevent overproduction. Other listers, like Isabella Lalonde, are crafting the designs that are setting the trends of the future. A year after graduating college, she launched Beepy Bella, a jewelry and accessory label that caught the attention of young stars like Olivia Rodrigo, Grimes, Bella Hadid, and Bad Bunny. “I remember Grimes messaging me on Instagram when I just started, saying ‘I love your work’—and it feels nice that people whose work you respect, also respect yours,” Lalonde says. After having her jewelry showcased in HBO Max’s Euphoria, she has since taken her L.A.-based business to retailers, including Nordstrom, SSENSE, and soon, Urban Outfitters.

Supermodel-turned-entrepreneur Kendall Jenner made the list with her first venture into the business of premium tequila. Forbes estimates that Jenner’s 818 Tequila made around \$25 million in sales in 2022.

Meanwhile, art adviser Kendra Walker is blazing a trail in her native state of Georgia. The 27-year-old founded Atlanta Art Week in 2022. “There’s great talent in Atlanta,” she says. “I think an art week within the city will create more long-term relationships with larger players throughout the world.”

Farther north, the state of Vermont appointed Tillie Walden its newest cartoon laureate, and at 26, she’s the youngest artist to receive the honor. Walden, who has published more than 10 books, often centering queer narratives, is also one of the youngest recipients of an Eisner Award, often referred to as the Oscars of the comics industry.

In Canada, Nia and Justice Faith Betty launched Révolutionnaire, an apparel company for dancers of color and have since sold out of collaborations in stores across Canada, the U.S., and Taiwan. The sisters champion social impact and activism, having worked with the nail polish brand Essie to create inclusive hues and raising \$1 million for volcano relief following the 2021 La Soufrière eruption in Saint Vincent.

While dividing his time between New York City and Los Angeles, celebrity stylist and author Andrew Gelwicks has a clientele that includes Sarah Jessica Parker, Catherine O’Hara, and Idina Menzel. And making impact around the globe is Gisela McDaniel: an indigenous Chamorro visual artist whose work has been shown in Germany, London, New York, India and at galleries such as Jeffrey Deitch, the Museum of Fine Arts Boston, and the Institute of Contemporary Art Miami.

This year’s Art & Style list was edited by Simone Melvin, Mckenna Leavens, and Osman Yerebakan

Collina Strada Takes On AI Fashion At New York Fashion Week

Rose Celestin, Contributor

Sep 18, 2023, 05:31pm EDT

Collina Strada - Front Row & Backstage - New York Fashion Week - September 2023

New York Fashion Week officially came to a close last week as the unveilings of many Spring/Summer 2024 designs from fashion staples to up-and-coming designers graced runways with celebrity sightings from Hollywood it-girls Blake Lively, Winnie Harlow, and Sofia Richie to Billboard artists Swae Lee, YG, and Ella Mai.

But apart from new spring trends ranging from pedal pushers, slouchy shoulders, cinched drop waists, and corset lacings, along with asymmetric necklines, pops of white and lilac, and avant-garde silhouettes, another notable, emerging trend that could potentially take the fashion world by storm is AI fashion.

While dystopian fears of AI's evolution across various industries have cast a slight shadow over the emerging technology's many use cases, Hillary Taymour's unveiling of her Collina Strada Spring/Summer 2024 collection at New York Fashion Week flaunted a more

"Everything sucks. We're all doomed. The world's on fire, but we're doing a fashion show because that's what we know how to do," Taymour told Vogue days before the show.

Collina Strada is a fashion brand centered on climate awareness, social awareness, change, and self expression. The eco-conscious fashion designer, Hillary Taymour, is headstrong on incorporating sustainability practices into her fashion line utilizing rose silk, an eco-friendly, biodegradable organic cellulose fiber made from rose bushes and stems, deadstock fabric, which turns unwanted landfill fabric into something useful, and recycled cotton.

But what happens when fashion meets at the intersection of technology? Disruptive innovation in the fashion industry.

According to McKinsey and Company, generative AI could add between \$150 billion to \$275 billion to the apparel, fashion, and luxury sectors' operating profits. This can in turn set a new precedence for creativity as unstructured data sets in the forms of raw text, images, and more can translate into realistic 3-D designs to boost productivity in the fashion realm.

Generative AI in the fashion industry can also be leveraged in certain go-to-market use cases such as supply chain and logistics, marketing, store operations, digital commerce, and consumer

experience to predict trends, automate and scale consumer segmentation, generate and test store layout plans, facilitate virtual styling and try-on assistance, and more.

Collina Strada - Runway - Spring/Summer 2024 New York Fashion Week

Taymour debuted her AI-influenced fashion designs during her “This Is Fine”, meme-inspired Collina Strada SS24 Soft Is Hard runway show on the rooftop of Brooklyn Grange, against the backdrop of “Why are we here, the earth’s on fire” lyrics sung by musician and collaborator cast in the show, Oyinda.

When Taymour and her team fed all of Collina Strada’s previous collections into an AI model, what yielded in return were photorealistic Collina Strada-esque aesthetics that—with the aided touch of human refinement—bred the looks for the Spring/Summer 2024 lineup, designed to “reconnect to the universal feminine energy during a time of political and social strife”, according to Taymour’s runway notes.

“Hillary is very inclusive—probably one of the most inclusive designers I know,” Oyinda told Vanity Fair during the Spring 2023 show last year.

Known for injecting fantasy into her looks, Taymour wanted to create a world and a show that was natural to her for the Spring/Summer 2024 collection, the fashion designer told CNN.

Using AI as a tool to “remix old Collina and drive it further,” getting AI to make the designs was the easy part. Taking the AI-generated sketches and constructing real-life outfits for production was a concerted effort Taymour’s design team managed to pull off over the course of several weeks despite the algorithm’s initial whimsical and phantasmic design ideas.

Although Collina Strada appears to be the first fashion brand to use generative AI to create physical Fashion Week looks, Robbie Barrat is an artist who foreshadowed AI-generated fashion a few years back, and may have pioneered the commercialization of AI fashion.

Having designed a Kanye West AI model that writes rap songs based on the rapper’s entire discography, Robbie’s passion for AI and neural networks led to a prestigious Nvidia internship

and a research stint straight out of high school at Stanford University's Center for Biomedical Informatics Research, where he also guest lectures Continuing Studies classes on artificial intelligence. The self-taught genius then lent his prodigious skillset to create Balenciaga AI, a futuristic Balenciaga fashion collection and runway lookbook in 2018.

"A question I get a lot about this is whether or not the AI is 'creatively' making these outfits, or if it's just mimicking Balenciaga. I don't think that the AI is being 'creative' when it comes up with this stuff, but I don't think that it's just mimicking Balenciaga either," Barrat said.

Trained on thousands of images from Balenciaga lookbooks and fashion shows, Barrat's Balenciaga AI model was able to come up with its own interpretation of what a Balenciaga outfit should look like.

"When humans are designing clothing, we know all about the nonvisual context our clothes have (like what bags are used for and why people carry them, why people wear coats, etc.). The network really doesn't understand or care about this stuff, so instead of a bag it might instead just generate a piece of cloth for the person to hold—or just generate a pair of pants with a big compartment built in because it doesn't understand that bags are separate from pants, since in all the images it sees they're always right next to each other. It also doesn't understand symmetry at all, but I really love the asymmetrical outfits. It's just, like, a totally alien perspective," Barrat further explained.

The first ever AI Fashion Week earlier this year showed the progress of AI fashion, where innovative designers came together to showcase their diverse range of designs and silhouettes made up of 15-30 different looks while incorporating the latest fashion trends into their AI collections.

Collina Strada's Spring/Summer 2024 collection further cements the use of AI as a potential innovative disrupter in the fashion industry's value chain. Apart from turning sketches into photorealistic designs for production in order to accelerate productivity, enhanced go-to-market capabilities can also foster cost-effective opportunities within the fashion industry in the long run.

AI in Fashion: The Good, the Bad, and the Controversy

BY SIMONA TOLCHEVA

PUBLISHED JUN 24, 2023

AI in fashion has the potential to make our lives easier. But it doesn't come without controversial drawbacks.

robot and human arms coming together

AI has made its way into a ton of creative industries. You can use it for art, song and text writing, and much more. And it seems like only a matter of time until it finds a place in every industry .But given the many issues people have rightfully pointed out regarding art and AI, is AI coming to fashion something to anticipate or dread? Is AI in fashion a good thing or one that warrants criticism? Let's explore.

AI in the Fashion Industry

Already, you can find AI has made its way into the fashion industry as you can use it to predict trends—with certain styles, colors, anything. AI can also help with design generation.

Fashion trends come and go much quicker nowadays than even a decade ago, so turning to AI to predict what's coming can give a leg up to fashion designers. You can easily test color and pattern combinations and base your next designs on what AI suggests.

AI can make a great personal stylist, too. Many retailers are exploring integrating AI in interactive smart mirrors on shop floors or changing rooms. Mirrors that will allow you to select outfit options and, with the help of AI, see how they fit on your body without putting on anything.

And as much as that sounds like a dream, the question remains, does AI deserve a place in fashion?

AI in Fashion: The Good

AI can benefit the fashion industry in certain aspects; for example, AI can make the process of shopping way easier.

If you can browse clothes, accessories, and everything else you need, without having to look around a store and try on anything physically, it's a game changer. Plus, it will take much less time and energy.

And, if you look to the future and hope for the best, AI can help streamline shopping even more. For example, you can fill a store with items of one sample size and keep it there for reference. You won't sell the things that are physically in the store but hold onto them, so shoppers can have something to look at and feel.

With the help of smart mirrors and AI, shoppers can try everything available at the location, see how they like it, and order it for pickup or delivery. That way, you don't fill the store with a mountain of the same items in different sizes and colors. This way shoppers don't get overwhelmed browsing and, what's more, the method is size-inclusive.

Everyone can walk into a store, and once they find something they like, order it in their size. And if we employ wishful thinking, maybe the items will get produced only after they've been ordered. That way, you'll avoid overproduction, won't add to landfills, and help the planet a bit in the fight against climate change.

Yes, you might have to wait for your item anywhere between a few days to weeks or even months. But it's ethical consumption in a capitalist society. Plus, the likelihood of you returning the item is less since you already know how it looks on you, and the AI has ordered the right size for you.

It's also beneficial for garment companies as it allows them not to waste money stocking every shop they sell at in all available sizes and colors for every item. And instead, have them shipped out once someone orders in a particular size, style, and color.

Even if the items are already made and simply stored elsewhere, like a warehouse, it still helps streamline the shopping process. So, looking at AI in that light, it could revolutionize the entire fashion industry.

AI in Fashion: The Bad

Modeling is among the areas in fashion with which the public tends to have many problems, even today after supposed years of progress. Among those issues is diversity.

There has been an outcry for diversity in modeling for a long time. People want to see diverse skin tones, shapes, sizes, and hair types. And not only on a runway.

Online shopping is a straightforward process, but if you wish to see how a garment or accessory fits, you're limited to seeing it only on the model shown on the website you're visiting. If that model differs from you, you must guess how the item would suit you.

And here's where diversity would improve the fashion industry. If AI easily and quickly generates all manner of models, brands can utilize considerably more people to showcase items. So, presumably, you'll be more likely to find models similar to you.

When you see yourself in the model, you're more likely to buy, so it does make sense why brands are starting to turn to AI to generate more diverse and inclusive models.

But many have expressed opinions that this is a way for companies to make their lives easier while saving money and face. After all, instead of hiring actual people and using them to help with diversity, they're using AI for computer-generated inclusivity.

Why hire and deal with people when you can buy AI-generated images of the people you would have otherwise used? And herein lies the controversy...

AI in Fashion: The Controversy

Though AI-generated models may sound grand in theory, the public seems to disagree about their practicality. AI models are branded as a tool to improve inclusivity and diversity, but it doesn't showcase actual people.

You're showcasing fictional individuals and don't know who's profiting from their image. So, while you believe you're supporting diversity or inclusivity by showcasing a specific type of person, in the end, you don't know who's profiting off the AI model you decided to use. So is this really the way forward?

Levi's AI Ad

Levi's found itself in hot water after including an AI model in one of its ad campaigns. The company turned to LaLaLand.ai—a digital studio that creates customized AI models—and Levi's was quick to point out that it will still utilize real people, too. And AI-generated models will only add to its roster and showcase a better representation of sizes, skin tones, and ages. But the brand won't invest in real human models to achieve that. Instead, it'd prefer to pay for AI to render them out.

Sinead Bovell and AI

Sinead Bovell is a model who's consistently spoken about AI making its way into our lives. AI is used in music production, it impacts healthcare and the game industry by introducing AI-generated games, you can create art with AI, and so much more. Given she's a model, she's shared a particular interest in AI in the fashion industry.

And her opinions have been featured on talk shows, publications, and more. She believes that although AI has certain benefits, the negatives outweigh the positives.

For one, AI may push her and other models out of the industry. Sure, AI can be good for models if they agree to license their image, for example. That way, they can work at multiple places at the same time.

But it's not certain whether AI-generating services would choose to pay for licensing images or just create new faces.

Is AI the Way Forward for Fashion?

As with most things, there's yet to be a clear answer. AI can undoubtedly help the fashion industry move forward and do better. But, sadly, in some aspects, it can do that at the expense of real people. And that itself leads to the question, will this be the new normal?

Are we to expect AI to become a part of our profession only to then push us out? Should we prepare to get replaced by AI? And how do you survive in an AI-forward space? Models partaking in the fashion industry are rushing to answer these questions for themselves today, but what about tomorrow?

(Tolcheva. 2023)

Chapter 5 -Film and Video

The incomparable cinema director, Guillermo del Terro, said, in 2023, “I am not afraid of artificial intelligence. I am afraid of human stupidity.”(Fu, 2023) With three Academy Awards, and unanimous critical acclaim, his point is vital to consider. That is because he has been using the state of the art media processing and generation technology earlier than any director and to a most natural affect for decades. His use of effects make his films more authentic through an organic dream state with which he imbues them. To clarify his statement: he means stupidity of company owners, investors, politicians and international bodies. He is not attacking anyone. He is telling them to get it right.

Justine Bateman took the lead in the industry’s union ecosystem, galvanizing solidarity and urgency before it was too late. Here is her interview with Tech Policy Press.

Justine Bateman on AI, Labor, and the Future of Entertainment

PUBLISHED BY TECH POLICY PRESS: JULY 23, 2023

Artificial intelligence will likely impact every type of job. But this summer, Hollywood actors and writers have raised substantial concerns about the ways in which generative AI systems may be used to replace aspects of their human craft. The Screen Actors Guild-American Federation of Television and Radio Artists (SAG-AFTRA) and the Writers Guild of America (WGA) are currently joined in a dual strike, hoping to make progress on a range of labor grievances with the studios and streaming companies that employ them.

Today's guest is [Justine Bateman], a writer, director, producer, author, and member of the Directors Guild of America (DGA), the WGA, and SAG-AFTRA. Bateman has been on both sides of the camera for much of her life, and has a particularly sharp perspective on how AI may change the entertainment industry, and why it matters to all workers that the unions are standing up on these issues now.

What follows is a lightly edited transcript of the discussion.

[Justine Bateman]:

My name's [Justine Bateman]. I am writer, director, producer, author. I am a member of the DGA, the WGA, and SAG. And if I found a way to belong to IATSE and Teamsters, I could have the entertainment union EGOT.

[Justin Hendrix]:

And I should point out also as well, someone who holds a degree in computer science and digital media management, which I suppose together with all of that experience makes you perhaps one of the best placed people to talk about the implications of AI and the current strike in the entertainment industry more generally. I was struck by your piece in Newsweek, your imagination around how you think AI will play out. And I'd love for you to perhaps share that with my listeners. What do you think's going to happen to the entertainment industry as generative AI and other forms of AI become more prominent?

[Justine Bateman]:

Well, for anyone who doesn't know, and I'm really simplifying it, what generative AI even means. Just picture a blender or a box. And on its own, it can't really do anything, it needs to be fed something. So generative AI in the arts you can imagine is if you want to produce a cool painting, you feed it as many paintings as possible, then you give the AI program a task and it'll spit out an amalgamation of what you fed it. Same thing with scripts, books, acting performances, entire films. So people can see how egregious the idea of generative AI is in the entertainment business. Because not only would you be replacing our work, but you'd be replacing our future work with an amalgamation of our past work.

Here's how I see things going, and if anybody wants to see demos of this, I encourage them to go to my Twitter account or my Instagram account. I put a lot of videos on there, and this is just consumer facing software, these demos, this isn't even the high end, more expensive option that I'm sure NVIDIA is offering many, many people right now.

Another interesting point for people to look at, and then I'll get into the progress of how I think it's going to progress or rather the steps I think this progression will take. I encourage people to look at this article. I think it was an adage about WPP, the ad agency conglomerate. They own Young & Rubicam and all these different ad agencies, something like Publicis Group. They own a bunch of ad agencies as well. And look at the relationship they have developed with NVIDIA, and now they're at a point where they're almost merging. And NVIDIA, as people probably know, they produce GPUs, which are essential for graphic processing. And I know they do a lot of other things, but that's the only thing that I have been focused on with this company.

And when they say things like we and excitedly in this article, Nvidia said, "We are going to be putting together the algorithm, the methodology, the structure, the infrastructure." Basically the entire setup for WPP, for their brands to send out individualized ads. Now, if I'm a brand at some point very soon, I look at WPP and I wonder why I'm paying them. I'm going to make a deal directly with NVIDIA because if they have this entire set it and forget it generative AI, it's more complicated than a algorithm people know. Then I'm not using a creative director, I'm not using

any copywriters. What do I need WPP for? And I think the same type of thing will happen in entertainment very soon.

So I think the process goes something like this. In entertainment, right now we have of course directors and writers and actors and cinematographers and a crew and the distribution and all of that. But what if we do generative AI, fed it in all of these films and then made available to viewers something that was customized to them, customized to their viewing habits. And they have many years of every individual's viewing habits. And let's say you watch films that are Hong Kong action films and you also watch a lot of nature documentaries. So they're going to put together nature documentaries/Hong Kong action films, et cetera.

So that's one thing, these customized pieces. Then they can also for an upcharge, ask every viewer to go to a particular establishment and go get themselves scanned and then they can put them in the film. Another thing, they'll ask you to upload a picture and which is all you need for face replacement. And they can replace, let's say you order up or I want to watch Star Wars tonight and put my face on Luke Skywalker's body, and I'm talking about things that are currently available.

And then for the future, and I don't know how close we are to something like this, but I think it's definitely on the very near horizon is feed in all the episodes of an old show, let's say M*A*S*H or Fran Drescher's show, The Nanny or something. And let's say they only did six seasons, they could generate a 7th season without any overhead, but the cost of the generative AI processing.

[Justin Hendrix]:

You've pointed out that not only will this technology affect actors and writers and directors, but also other folks in the industry who are behind the camera, everybody from the caterer on through to the grip and the lighting folks. It's not hard to imagine the footage and the content that's generated by generative systems being put into virtual environments, maybe in some game engine and manipulated in post-production without terribly many humans being involved at all.

[Justine Bateman]:

And I think that option is more enticing to the studios and streamers right now than anything else. I don't know what Netflix, what did they spend on making films and series last year? I don't know what the number is. Let's say it was \$4 billion and let's say they made \$7 billion. I don't know what the numbers are, I'm just making things up. What if you could make \$7 billion and only have spent \$1 billion or half a billion? What if you could cut that overhead?

It's pretty difficult to cut it if you're not using generative AI. But if you are, then all those people you've just mentioned, when you hire people in the entertainment business, if they're part of a union, you not only have to pay for their work that day, but you also have to pay fringes, which is their pension and health. You have to feed them, you have to put them up if you're on location, there are all these other expenses that are beyond just that salary. And if you could cut that out, then your profit margin would be incredible. And I think that's the motivation in all the sectors that are adopting some form of generative AI, whether it's just text or speech or full-blown video.

[Justin Hendrix]:

So what you're saying is that to some extent, you imagine the executives who are running these media and film companies to some extent, the perfect price of content is zero?

[Justine Bateman]:

Yes. And it's never been closer to doable.

[Justin Hendrix]:

So I want to talk a little bit about this moment. You have pointed out in prior interviews that you know think that this is perhaps the last moment of leverage that the industry will have. Why is that? What makes this moment such a last stand in your mind?

[Justine Bateman]:

Well, you can see right now with everyone going on strike, anything that was in production stopped. People just left the locations wherever they were shooting, the actors did. With writing, it's harder to stop things on a dime because there are scripts that the studios or streamers had

already purchased that they could go ahead and shoot. The same thing would've happened if directors went on strike. Everything would've just stopped on a dime. Now, if you're making things with generative AI, nothing stops. So I feel, and this is also based on what the AMPTP lawyers, and for anyone who doesn't know, that's the association of all these streamers and studios, where they bargain together. Based on what they asked for from the writers and the actors in terms of generative AI, indicates to me that they are further along than anybody knows.

I'll give you comparison. I was on the negotiating committee for the Screen Actors Guild and on the board in '07, '08. So during that negotiation when we were talking about made for new media, that was the new area. YouTube had a upload restriction of three minutes, that's how long ago this was technologically. And we said, "Okay, well what about when you distribute stuff online?" And they're going, "Made for new media, that's so new. We don't even know if one can make any money over there." And I was already working in digital media at that time as a writer and producer. And so I said to the negotiating committee, I said, "This is bullshit. This is just another way to distribute things. Trust me, you've got to get real estate in here. We have to have rates in here, we have to have residuals. This is just like when TV began and it's just another way to distribute the same work."

And we got some, but not enough. And still to today, still today rather, and still today, there are gains that we need to make in that area, all the unions do, to catch up to where they have... Our increases with made for new media residuals have not at all matched the degree to which these studios and streamers have succeeded with made for new media, what it used to be called. So contrast that with now, the asks. They didn't even say, "Why would we talk about AI? We don't know if that's going to be useful at all." When they said to the Writers Guild, we are not talking about it, I immediately knew that they had already fed in whatever scripts they could get ahold of into some kind of LLM.

When they said to the actors, "Okay, we will give you consent and some sort of compensation if it's a direct digital double of you, but we'll only give you a little bit of compensation and no consent if we're just feeding into a generative AI model and making sort of Frankenstein where

you can order up like, 'I want a character that's a little bit of Brad Pitt and a little bit of George Clooney who dances like Fred Astaire, these kind of Frankenstein orders.'" That to me was very telling. I was like, "There's no way. We can't allow to not have any consent." You can just use all of the work from the past and not ask anybody if they're okay with being replaced by an amalgamation of their work? It's crazy. And I think SAG gives up that, you handed them the keys to the house because then no longer to the point you made earlier. Then you no longer need any of the crew members on a set.

Listen, I think all this is going to happen anyway, but at least we can set some sort of example for how other industries can push back. And so in that sense, this, I believe 2023 is the last time that any of these unions will have leverage on their own. I think in 2026, if DGA WGA and SAG banded together, I think that might be the only way we have leverage. And maybe somehow we can get the IATSE and Teamsters cycle to match up with ours. And that would mean extensions and stuff. And I don't know how possible that is. But I'm just saying in three years, I think they'll be so far along with AI only films that I don't know that they'll care if we go on strike.

[Justin Hendrix]:

Is there a difference between the interests of the WGA and the Screen Actors Guild and other unions on this in terms of what they want?

[Justine Bateman]:

There's some things that are very similar. Generally speaking, it's extremely similar. But of course the writers have some asks that are unique to them. TV series writers used to be engaged for the entirety of a series, and the series also used to be 22 episodes or around there instead of now eight. And now they've reduced the writer's involvement in many cases to get in a room, four weeks, eight weeks, something, write all the scripts for the entire season and then go away. The showrunners probably had experience on a set doing a show, but none of these other writers and that's your farm. That's how you build new showrunners. So that's one thing that is an extreme problem.

And then I think one thing we all have in common definitely is a concern with AI and then streaming residuals. The reason residuals were established many, many years ago, I'm not going to remember the story exactly, but I believe it was musicians on radio. I think they were like two shows a day or something that they were contracted to do and they would pay them for the two shows. And then they said, "Well, listen, we're going to record the first show and just play the second show. You guys don't have to come in." And the musicians said, "Well, you just took a session away from us. You have to pay us for that too. You can't just replace us with and not compensate the loss of the second session."

And so in that sense, if streamers just want to and basic cable's a very low residual structure as well. But in particular streamers, you want to have all this. You're just a website. If you don't have all of this material, you're just a website with no reason for anyone to come to it. So if you want to have all this material on there that's just available endlessly, you can see how that displaces anybody's additional employment, like those radio musicians or those musicians for the radio, the point they were making.

[Justin Hendrix]:

Is there some context in which you would like to take advantage of AI or can you imagine actors, writers, others in the industry who are currently advocating for their own interests, seeing some of the benefit of artificial intelligence? Have you been presented with the opportunity to license your avatar in contracts, for instance?

[Justine Bateman]:

It's something I am 100% disinterested in. So for me as a filmmaker, I'm going in the complete opposite direction. I want to make new things that have a deeper emotional impact to the viewer than any work I've done before. Or ideally, what if I could make something that had a greater emotional impact than any film before? I don't even know if that's possible, but if you have that as a goal, you can't use AI for anything. It's generative, automatic imitation. So all it is going to

regurgitate the past. I would never be making anything new, and I would also be stealing from myself of my own enjoyment to do anything, to make film.

I love writing and directing. Why would I give that away to somebody? So for me, it's not for me, but the job of these unions is to set a floor, set a default, and the default for actors and writers and directors too. I'd like to see directors' past work protected because now they can feed it all in and just say, "Well, I want something that's in the style of PT Anderson, or "I want something that's in the style of Alfred Hitchcock or something like that." Which I think we have a responsibility to not necessarily older, but to directors' past work. We have a responsibility to all of those directors.

So if the default for, say, Writers Guild or Screen Actors Guild is no, then that takes the pressure off them. That's what a union is for. A union is supposed to establish the floor. What's the default? The default can't be yes for things like this where then it's up to the individual to say, "No, I don't want to do this." It's up to the union to say, "This is not allowed without compensation and consent." So that means they would have to go to the actor or go to the writer and say, "Do you consent to this? And what would you like to be paid?"

Particularly for actors, because they can be on location in say, Europe, on a film. The second AD knocks on their trailer asking them to come out and we're ready for you on the set to shoot a scene. You're in the mindset of doing this creative work, of playing this character. And then he hands you a document and says, "You need to sign this first." And you have to shift gears and say, "Well, wait a minute, what is this?" And they say, "Oh, well, we're going to do a scan of you later." And you ask, "What's that for?" And he says, "I'm not really sure, but you need to sign this. They say, you need to sign this before you come on the set." And then you try to get your agent on the line, but it's the middle of the night because you're in a different time zone. So middle of the night for the agent.

And so the actor themselves is faced with signing away their likeness and then continuing on to go to set or holding up the entire day. It's not fair to put any actor in that kind of position, which is why the default has to be no, and it has to be compensation consultation. And I think personally, you really got to spell out exactly what you're going to be doing with it, because it's also for an

actor... For a writer, it's horrible to take their material and do whatever you want with it. But a writer can write another script, which doesn't make it any less horrible, but for an actor, it's criminal.

[Justin Hendrix]:

It's your body.

[Justine Bateman]:

It's creative identity theft. We're going to see it in the election coming up. There's going to be so many things going on with generative AI in other areas of our lives that are really going to toss people's sense of reality.

[Justin Hendrix]:

I want to ask you about that, perhaps. Is there a message that you might have for policymakers who are considering AI more broadly, considering generative AI, perhaps specifically that perhaps you would give beyond the interest just of the entertainment and creative industries. What do you think policymakers should be doing around this technology?

[Justine Bateman]:

I think policymakers should get off their asses right now and do something about it right now. Europe didn't have a problem doing something immediately. I think in back in April, they started introducing legislation, honestly. I think it seems to me the problem might be that these enormous tech companies who are now over here part of the AMPTP, think about that. We're not just negotiating against the biggest companies in Hollywood. We're negotiating against the biggest companies in the world, and these extremely large companies have contributed to these political campaigns. And I think unless there's a politician that wants that gravy train to stop, they'll continue to drag their feet like they're doing right now.

I think there was a AI, they posted this, I forgot what it was called, but they had all these AI...

There was somewhat of a hearing. This was a big video presentation and they had all these people

come and talk and tell them more about what's coming and all of this. And I'm sure most of them talking about how exciting it is. And they came out of there, shrugging their shoulders. Quote after quote after quote in this article was, "Gee, I'm not really sure what we can do here. We don't want to stop innovation." They know exactly what they need to do. Exactly. Really any politician who said that, look at who their contributors are. And I bet you will find these big tech companies on there.

[Justin Hendrix]:

Perhaps these days, harder to find a politician that is not receiving funds in some way or other from one of the large tech firms.

[Justine Bateman]:

They can thank Citizens United and all the Supreme Court justices that let that go through where corporations could be seen as people and contributed in kind. I think that was the worst thing they could have done to U.S. politics.

[Justin Hendrix]:

I'll ask you maybe just two last questions. One is, you have pointed out how important the timing of the strike and perhaps the emergence of things like ChatGPT, just November last year, just a few months ago, really. This convergence. Do you think AI would be as big an issue as it is in this particular moment if it weren't for ChatGPT, if it weren't for OpenAI's release of that product late last year?

[Justine Bateman]:

Yeah, you bring up a good point. So our negotiation cycle for the entertainment unions is three years. I'm pretty sure that's true for all of them. I could be wrong. Definitely for WGA, SAG and DGA. So here comes our cycle the summer of 2023 or spring/summer of 2023. I started posting about the threat to the entertainment business in about February. So you say this became available to the public in November, and I'm thinking, what if it had become available to the public instead

of November 2022, November of 2023? And what if WGA, DGA, and SAG had negotiated? It just would've seemed like another round of negotiations. Not much has changed. Try and get more for streaming residuals and things like that, but basically go the same route as maybe the 2020 negotiations or the one before that.

And then they had released access to these generative AI models to the public, and we would've just had to sit there for three years and just watch it disappear. And we would've never had leverage again. So I don't know why it worked out that way, but it's pretty lucky. Because I had a Bloomberg journalist make a good point. She said, "It's great in a way that it's coming after the entertainment business first because you guys have really strong unions and people are also interested in your business. And so people are watching, people are listening." And maybe we can become a template for, or at least a model for how to push back and not just lay down.

And it's people that are like, "Oh, Bateman, you got to just go with it. It's the future. It's what's happening." And I'm like, "Give me a break." They're like, "It's out there. It's out there. There's nothing you can do about it." Yeah, I know there's nothing I do about it, but fentanyl's out there too. But I'm not going to go snort it because it's available and it's not the future. It's an absolute 100% regression. It is the opposite of the future. It is the opposite of innovation, it is the opposite of the new. All it is regurgitating everything that is available online or otherwise. You could feed it in anything. I think it's the absolute wrong direction for society and in particular the arts.

I don't think that I can stop it. I have been attempting to since February, just tell the rest of the people on the beach that the tsunami's coming, because I really love everyone who's called to this business, and I don't want them to be caught unawares because the AI infiltration into our society and into various businesses is insidious. It's something that will be brought in by whoever your boss is and set up underneath within the infrastructure of that company. And as soon as it's set up and in place, they'll say bye-bye to you. It's not going to be, "Hey, we're thinking about bringing this in." And that is the problem the AMPTP has right now is that they don't have it completely built up yet. They would've loved it if this had come out in November of 2023. If I was the AMPTP, I would've agreed to everything and then in the background, made my all AI films and

then 2026 comes around and I just say, “We’re not going to be a signatory anymore. We don’t need to negotiate with any of you.”

So they’ll probably do that anyway. I think this is really going to burn the entire business down. I’m not meaning that to sound like a dramatic idea. I really do think we will see the end of the entertainment business, of this 100 year old entertainment business. But those who are called to this, if they just hold on, I believe on the other side, viewers will, after they are done with the novelty of it and the novelty of seeing themselves in all of it, I think they’ll feel like the end of *Supersize Me*. They’ll start feeling sick. That’s where Morgan Spurlock ate McDonald’s every day for 30 days. Sounds great at first and then not so good.

And I think then they will start rejecting AI and they will start rejecting any “content” that seems like AI. And they will also be distrustful of anything they’re seeing or reading because it’ll be so much infiltration of AI. And I think they will want to see and hear. They will want music and films and series that are really raw, really human, really obviously made by humans. And I think that will lead to a new genre in the arts, and we haven’t had one of much significance, I should say, since the 90s because tech has been center stage for like 15 years solid. And it needs to step aside and of other things in life, be in the middle here.

[Justin Hendrix]:

Any predictions on how long the strike will go on?

[Justine Bateman]:

A prediction? I would predict a month. If I’m the AMPTP, I let them whoop it up for two weeks in this heat, and then I give them another two weeks to start feeling the heat and then I start. I don’t know, it’s interesting. They’ve put themselves in a bad position because they’ve got a lot of films that they need to release right now. A lot of actors that they need to help promote these films. So they’re in a bit of a negotiating bind, I think. So that’s just a very weak prediction. That’s not me saying, I think this is the way it’s going to go. I’m just throwing out a 30-day thing into the just in prediction pot.

[Justin Hendrix]:

Well, perhaps I can bring you back on in a couple or three years time at that next cycle. We'll see what the impact of AI was in this moment. And perhaps if some of your predictions about this tide that's about to wash over the industry come true, hopefully your efforts to educate and engage with folks will perhaps improve matters. But I thank you so much for joining me today.

[Justine Bateman]:

My pleasure.

Justin Hendrix is CEO and Editor of Tech Policy Press, a new nonprofit media venture concerned with the intersection of technology and democracy. Previously, he was Executive Director of NYC Media Lab. He spent over a decade at The Economist in roles including Vice President, Business Development & Innovation. He is an associate research scientist and adjunct professor at NYU Tandon School of Engineering. Opinions expressed here are his own.

Justine Bateman expressed concern, on Twitter, over a number of issues in the Screen Actors' Guild [contract](#) that emerged in November. (Actually, it had been too late.)

#SAG Actors, I want to make you aware of some of the language in the #AI portion of the tentative SAG agreement.

Though SAG leadership made much effort to protect members in AI, there are many issues you should look at. I have saved the most serious issue for the end.

If a "digital double" is made of you during a film, they must get your consent and inform you of their intentions for its use, EXCEPT "when the photography or sound track remains substantially as scripted, performed and/or recorded."

This is going to be left up to the studios/streamers' interpretation. And so, any subtlety regarding how you chose to look or move for the character during the shoot could potentially be changed. Your hair, your clothes, your make-up, etc.

Also, your physical placement in a scene can be changed, like your nearness or distance from another character, or even moving you from the front seat of a car "to the back seat of the car." This suggests not much agency on your part to control your character or performance.

Under "(Digital Double) Use Other Than in the Motion Picture for Which the Performer Was Employed," it says that "No additional compensation shall be required for use of an Employment-Based Digital Replica that was created in connection with employment of a performer who was employed under Schedule F." It appears that if you were paid Schedule F for the first film, you don't get paid for the sequels, where they're just using your digital double instead of you. I suggest members get sharp clarity on this.

If a "digital double" was made of you in a separate manner (on another film or privately made by you), it's referred to as an "independently created digital replica" (ICDR). There is no minimum compensation listed for studios/streamers to use an ICDR of you in any film they want; only consent. You will apparently need to negotiate any compensation on your own.

Neither consent nor compensation is necessary to use your "digital double" if the project is "comment, criticism, scholarship, satire or parody, a docudrama, or historical or biographical work." So, you could find yourself in a project you never consented to doing things you never were informed of, for no compensation at all. This is the "First Amendment" argument the #GAI tech companies are fond of trotting out.

Another consent exemption is granted to "adjusting lip and/or other facial or body movement and/or the voice of the performer to a foreign language, or for purposes of changes to dialogue or photography necessary for license or sale to a particular market."

The substitution of swear words is not new, but that your "body movements" would be changed suggests you'll be used like a type of rag doll in post-production.

There are still a few concerns with the Background Performers' details, but there's one that stands out as especially sad. "If the Producer uses a background actor's Background Actor Digital Replica in the role of a principal performer, the background actor shall be paid the minimum rate

for a performer... had (they) performed those scene(s) in person." So, if an extra is "bumped up" to a principal cast member, they never get to experience that position on a set. But you get a check after the fact.

And the most serious issue of them all is the inclusion in the agreement of "Synthetic Performers," or "AI Objects," resembling humans. This gives the studios/streamers a green-light to use human-looking AI Objects instead of hiring a human actor.

It's one thing to use GAI to make a King Kong or a flying serpent (though this displaces many VFX/CGI artists), it is another thing to have an AI Object play a human character instead of a real actor. To me, this inclusion is an anathema to a union contract at all.

This is akin to SAG giving a thumbs-up for studios/streamers using non-union actors. This would be like the @Teamsters putting in their contract that it's A-OK for the employer to utilize self-driving trucks instead of them.

I find it baffling that a union representing human actors would give approval of those same actors being replaced by an AI Object. And don't forget, those AI Objects are a mash-up of all actors' past performances, adding insult to injury.

Bottomline, we are in for a very unpleasant era for actors and crew. The use of "digital doubles" alone will reduce the number of available jobs, because bigger name actors will have the opportunity to double or triple-book themselves on multiple projects at once.

The use of these "digital doubles" will most likely preclude the need of a set or the use of many @IATSE crew and @Teamster drivers.

Audition odds will change. Winning an audition could become very difficult, because you will not just be competing with the available actors who are your type, but you will now compete with every actor, dead or alive, who has made their "digital double" available for rent in a range of ages to suit the character. You also will be in competition with an infinite number of AI Objects that the studios/streamers can freely use. And a whole cast of AI Objects instead of human actors eliminates the need for a set or any crew at all.

You are a complex & remarkable human. Don't let the CEOs convince you otherwise. Seek out filmmakers & showrunners who value your basic worth & committed to human workers on their projects. These are the ones who will make work that matters. We're going to be OK. Just hold on.

There is an additional protection I find missing from the #SAG #AI deal: protection for past work against those actors' performances being face-replaced or "re-skinned" using a "digital double" or a human-style AI object.

And has been highlighting the allowance for SAG itself to give permission for use of your likeness after death (if the heirs can't be found after "best efforts," which means minimal or no effort). This is a very big issue on its own and if any compensation is give to SAG for its use, that effectively makes SAG a talent agent, which is an enormous conflict of interest and does not incentivize the union to bargain properly in the future for live human actors.

In fact, #SAG demanding "notification" and "consideration" (presumably financial) from studios/streamers using human-like #AI objects, effectively makes them the talent representatives of these "Synthetic Performers." Conflict of interest.

Deep Fakes

Deep fakes are potentially extremely harmful with regards to misinformation and revenge porn. Blockchain will be part of the solution to dangerous deep fakes, in conjunction with personality rights protections. 85% of the victims of deep fake revenge porn are women. The UN Sustainable Development Goal #5: Equal Rights for Women will provide impetus for international harmonization respecting these vital legal protections. (Tyagi, 2022)

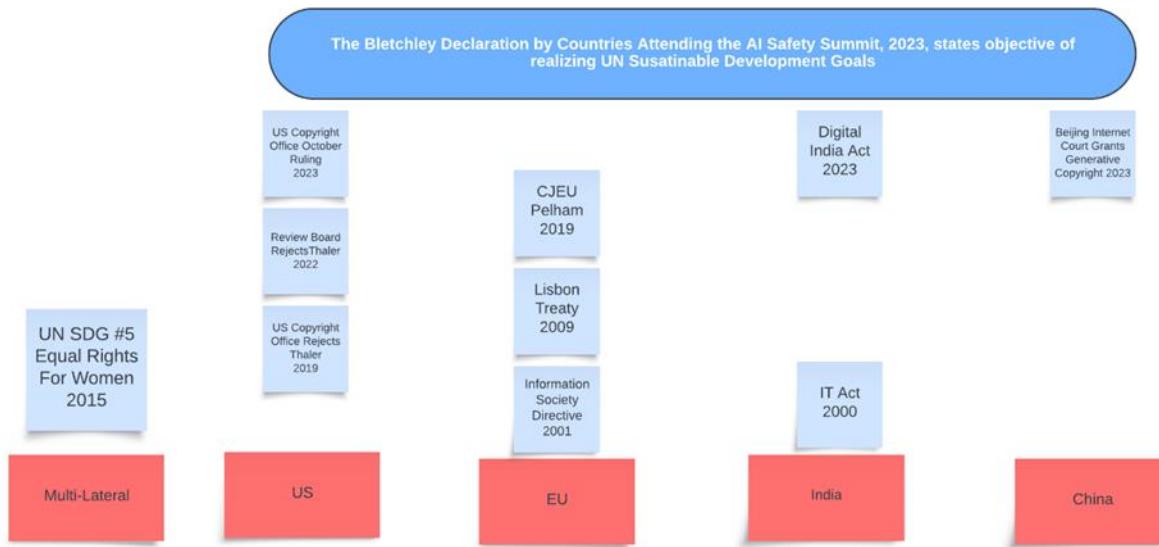


Fig. 3: Deep Fake Personality Rights and The Brussels Effect

Section II -Non-Creative Professional Sectors

Chapter 6 – The Gig Economy

[Chen et al.](#) (2023) conducted a study of AI's effect on labor alienation. They used a questionnaire of 225 temporary and migrant workers, to inform the principles of new labor policy and laws. Their introduction is a comprehensive guide to the emergent relationships, effects, and lag times for AI affecting low, middle and high skilled workers.

Labor Ethics Study on the Employment Substitution Effect of AI - Based on the Perspective of Changes in Logistics Work

Introduction

While some employees have a neutral attitude towards the changes in employment modes brought about by AI (Brougham & Haar, 2018), the increasing use of AI has led to more people being employed in informal, unstable, and non-standard modes, which has weakened the collective power of employees (Boeri *et al.*, 2020). Unlike past automation technologies, the new generation of AI is replacing mental labor and changing the division of labor in unique ways, with differentiating features in the value structure of human capital (Atack *et al.*, 2019). AI has a strong "leader goose" effect with a high spillover driving force, as noted in a speech by Xi Jinping at the Ninth Collective Study of the Political Bureau of the CPC Central Committee

in 2018. Due to the weakening of external ability, low-skilled workers have limited autonomy in terms of the time and spatial range of skill improvement, resulting in slower skill improvement than the speed of AI substitution.

In recent years, some industrialized countries have faced the dilemma of declining labor share and slow growth in labor productivity despite the widespread use of AI (Autor, 2019). There is an increasingly deep contradiction between sluggish growth in labor productivity and changing employment demands. Using the wrong type of AI can lead to misallocation of resources by markets, enterprises, and individuals towards labor-intensive tasks, exacerbating labor substitution and causing serious unemployment, weak growth, and income inequality.

The degree to which workers are substituted by AI depends on the match between the types of daily tasks and skills. Workers with high skills who engage in complex work tasks are less likely to be replaced by AI because they possess strong cognitive skills. However, in the process of substitution, high-skilled workers have exacerbated the ripple effect of replacing low-skilled workers. This contradiction leads to significant substitution of AI for medium-skilled labor, affecting long-term employment equilibrium (Acemoglu & Restrepo, 2018b).

According to Acemoglu *et al.* (2021), workers highly involved in tasks with AI experience negative impacts on their employment and wages. The displacement effect primarily involves the substitution of programmatic work and some non-programmatic work, such as transportation, driving, and image diagnosis, where AI excels (Autor, 2015; Liu Xiangli, 2020).

As AI diffuses and penetrates different regions, sectors, industries, and departments of the national economy, the degree of skill premium becomes increasingly prominent. The demand for human capital shifts from medium and low skills to high skills, resulting in employment polarization and exacerbation of income inequality (Acemoglu & Restrepo, 2019).

Low-educated blue-collar workers, medium-skilled blue-collar workers, and white-collar administrative positions are gradually disappearing, with low-educated men facing a sharp decline in employment prospects (Binder & Bound, 2019). Low-skilled workers have lost core skills that cannot compete with technological premium (Agrawal *et al.*, 2019), and technological change has narrowed their scope of work (Acemoglu & Restrepo, 2020), leading to their replacement by AI in the fast automation process (Autor, 2019).

Under the labor selection mechanism, high-skilled workers are increasingly inclined to work in high-wage industry sectors and are more likely to form partnerships with each other, reinforcing the entry barriers to high-wage industries and isolating medium and low-skilled workers (Song

et al., 2019). Income inequality indirectly increases the relative income of capital owners, leading to an increasingly severe class divide (Acemoglu & Restrepo, 2020a).

The root cause of this issue is that AI weakens the ability of low-skilled workers to obtain subsistence materials from labor products, while strengthening the ability of high-skilled workers to obtain subsistence materials from capital products. *Through the gap in human capital accumulation, income inequality is indirectly exacerbated.*

Labor with scarce skills and low costs will bear the consequences of downward economic and organizational forces, and the chances of returning to formal work organizations are slim (Fleming, 2019). Although the degree of technology application will slow down the spread of the substitution effect (Naude, 2020), labor demand will be stifled by the inability to fully benefit from the creation effect, which cannot generate a strong creation effect to resist the labor demand decline caused by the substitution effect (Acemoglu & Restrepo, 2018a; Clifton *et al.*, 2020).

The job-creation effect takes a long time to materialize, and the creation effect of AI is unlikely to offset the negative pressure on labor demand, employment, and wages caused by the substitution effect in the short term (Lane & Martin, 2021).

...

Data Analyses

...

We selected four indicators to measure AI because it can optimize workflow and improve work efficiency, resulting in a shorter time to complete work units. AI technology can automate some tedious, repetitive, and monotonous tasks, while the remaining work cannot be completed by AI and can only be completed by employees. The inability to withstand new workflow is a result of AI. The improvement in platform feedback speed is due to the fact that AI can quickly process and analyze large amounts of data, providing faster and more accurate responses and feedback. Flexible employment personnel are characterized by decentralization, making it difficult to form a strong organization to protect their labor rights and interests. They also have a relatively low legal status.

Companies with high labor union organization rates can better meet the interests and needs of employees, safeguard their rights, and enhance their enthusiasm and work motivation. Similarly, the signing rate of labor contracts is an important indicator of whether a company and employees maintain a good cooperative relationship, ensuring the legitimate rights and interests

of both parties, reducing labor disputes, and enhancing mutual trust and cooperation.

A high turnover rate indicates that the company has a relatively high employee turnover rate, which may lead to employee dissatisfaction and instability, and job migration may have a certain impact on the career future of employees. High absenteeism rates may reflect a decrease in employees' work attitude and motivation, or an increase in absenteeism rates due to psychological factors such as anxiety and unease during job transfer.

The development of AI and other new technologies has brought about employment shocks, deteriorating labor relations, and increased willingness to relocate jobs. In the context of more intense economic competition, enterprises pay more attention to efficiency and cost control, and tend to adopt more flexible labor allocation methods. The problems brought by these new labor allocation methods mainly focus on labor ethics and morality. Enterprises may take unfair or unreasonable measures to reduce costs and increase efficiency, such as increasing workload, reducing wages, and strengthening rewards and punishments. These measures may lead to a decrease in employee satisfaction, a deterioration in labor relations, and a loss of labor ethics. The distorted relationship between humans and labor has turned labor into a form of oppression. And torture rather than self-actualization and creation, leading to the destruction and loss of workers' personality shaping, and a lack of space for creativity and self-realization. This phenomenon exacerbates the alienation of labor among workers, leading to the loss of labor ethics.

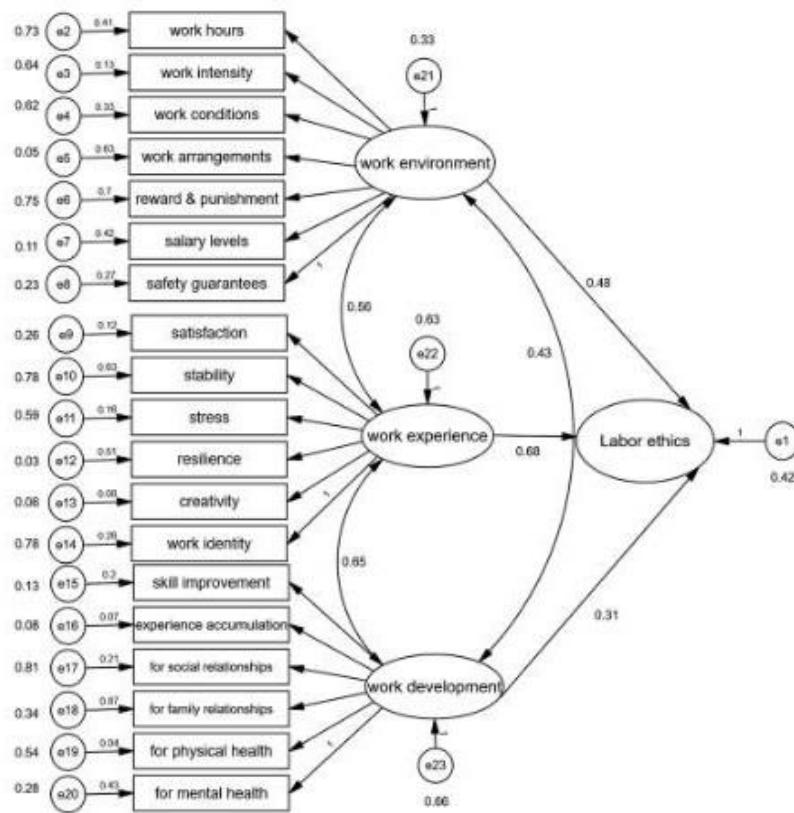
Result Analysis

According to Fig. [4]., the influence of work environment on labor ethics is 0.48, while the impact of work development is relatively weak at 0.31. The most significant impact is on the accumulation of work experience, with a score of 0.68.

As shown in Fig. [5]., AI impacts labor ethics by influencing labor relations and job migration.

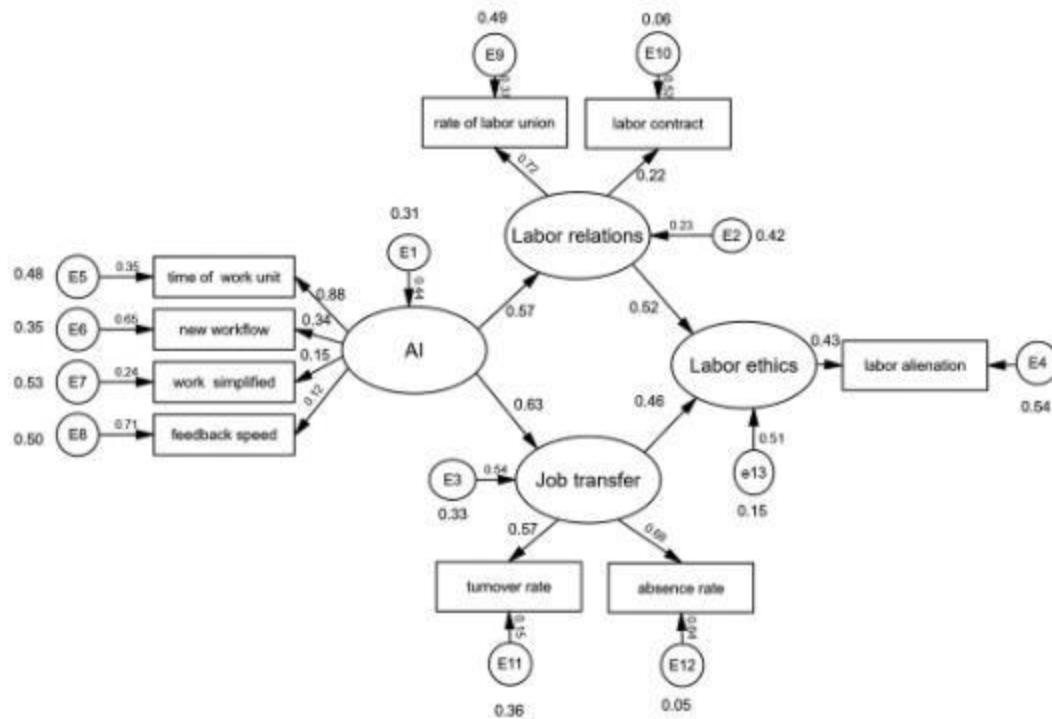
The impact of AI on job migration is 0.63, while its impact on labor relations is 0.57.

Additionally, AI has an impact of 0.46 on job migration and 0.52 on labor ethics.



Determinants of Labor Ethics

Fig. 4



The Structural Equation of AI on Labor Ethics

Fig. 5

Discussion

A labor ethics framework that is humane and respectful should consider factors such as AI, labor relations, and job transfer. Such a framework can promote individual workers' self-realization and achieve win-win cooperation among employees, enterprises, and society. Therefore, enterprises should manage and optimize their work environment, experience, and development comprehensively. They should provide employees with a work ethic and promote the stability and long-term development of the labor market.

As AI technology continues to mature, diverse forms of labor will become the norm. Regulating these forms of labor in labor law has become a challenging problem in both theoretical and practical fields. To address this issue, labor law must respond to new forms of labor

employment and adjust the impact of technological progress on labor law and labor relations in a timely manner.

Conclusions

With the development and challenges posed by AI, there is a need to reform existing labor laws. This includes clarifying the legal attributes and status of flexible employment personnel in non-typical labor relationships, determining the scope of application of non-typical labor, and defining the legal responsibilities of labor subjects. Additionally, there is a need to study the possibility of separating work related injury insurance from labor relationships. To address new occupational risks brought about by human-machine collaborative labor and remote labor, it is important to include flexible employees in the coverage of work-related injury insurance.

To balance the employment flexibility of employers and the employment stability of workers, specific labor standards need to be adjusted reasonably, and mandatory and autonomous norms should be reconstructed. This will achieve the recognition and compatibility of flexible employment forms in labor law. In the era of AI, it is important to shift focus from protecting standard labor relations to protecting atypical labor relations. The risk of employment instability caused by market-oriented mechanisms and flexible employment should be shared by society rather than borne solely by enterprises.

The gig economy has accelerated the dispersion of social power, leading to new types of labor capital conflicts under digital exploitation and triggering labor migration. It is important to guide migrant workers to establish a work ethic in the digital age, enhance the dignity of migrant work, optimize relationships between various parties in the migrant economy, and promote the healthy development of the migrant economy.

The real value of this study and the structural equation model will come from comparing it with subsequent replications in subsequent years and cities, including various countries. The values of AI, Labor Relations, Job Transfer, Labor Ethics, and Labor Alienation and the other variables will be more interpretable with geographic and time datasets for comparison. If the authors can team up with their international colleagues for subsequent studies, significant new indicators and interpretations could emerge.

Chapter 7 - Utilities

The employment in the utilities sector was projected to contract by 0.3% by 2033 in a U.S. Department of Labor 2023 report. The prospect of increased productivity due to AI, in the literature, is less significant than the impact on

exploration, safety, sustainability and net zero carbon. This critical review of AI in the energy sector in Vietnam gives a perspective of a developing country but it is also objective in a global sense. ([Pham](#), 2024)

Artificial intelligence (AI) development in the Vietnam's energy and economic systems: A critical review

...

On a global scale, in the energy sector researchers study many aspects of AI technologies although publication numbers are still limited compared to other sectors. The focus of research includes prediction, optimisation and control of thermal energy storage systems (Olabi *et al.*, 2023), applications of intelligent cyber-physical systems to decarbonise greenhouse gas emissions by improving energy efficiency (Inderwildi *et al.*, 2020), management of energy systems (Li *et al.*, 2022), management of industrial building energy saving (Zhao, 2023) and analysis of energy poverty (Bienvenido-Huertas *et al.*, 2023). AI technologies have also been considered key factors to accelerate the energy transition to low and zero emissions resources (Xie *et al.*, 2022). AI technologies applied to decarbonise greenhouse gas emissions levels, particularly in residential buildings, appliances and electricity supply, are also encouraged to help countries achieve net zero emissions targets at modest cost (Xiang *et al.*, 2023; Yan *et al.*, 2023). On the demand side, buildings show the most significant potential in cost-effective emission reduction and with support of AI and other technologies, such a sector can achieve carbon neutrality (Zou *et al.*, 2023). There are also many other aspects of the energy sector that have been studied, such as estimates of solar air heaters energy performance (Das *et al.*, 2023), parameter optimisation by AI in renewable energy-driven desalination systems (He *et al.*, 2022), the efficiency of solar still production (Mashaly and Alazba, 2017), energy sustainability in the internet of things (Charef *et al.*, 2023) and the identification of challenges and perspectives of AI-powered large-scale renewable integration in multi-energy systems (Liu *et al.*, 2022).

AI-related studies in the energy sectors in Vietnam are, however, scarce. Most AI-related studies focus on forecasting energy output and consumption. For example, Quang *et al.* (2021) and Nguyen *et al.* (2021) examine how AI technologies can be used to forecast the output of large-

scale power plants. A few studies also investigate the utilisation of AI technologies to predict energy consumption in residential buildings (Tran *et al.*, 2020a; Ngo *et al.*, 2022). Studies of AI in mining sectors are also scarce in Vietnam – those that are available examine the prediction of slope failure in open-pit mines (Bui *et al.*, 2020), analysis of slope stability (Luo *et al.*, 2021) and prediction of blast-induced air overpressure (Nguyen and Bui, 2019). There are also AI studies relating to the construction sector which identify the strength of cement-stabilised soil (Ngo *et al.*, 2021) or predict soil shear strength for road construction (Tien Bui *et al.*, 2019)...

Tanver *et al.* (2023) explain the means by which AI will optimize the energy sector and what the smart grid will look like.

Energetics Systems and artificial intelligence: Applications of industry 4.0

...2.1. The use of artificial intelligence in energy generation

Worldwide, the energy market faces rising growth, productivity, changing demand and supply trends, and the lack of technology needed to handle it effectively. In emerging market countries, these challenges are more acute. Efficiency problems are especially problematic, since the prevalence of informal power grid connections means that a large amount of power is neither measured nor counted, resulting in losses and greater CO2 emissions, since consumers are not encouraged to use their energy efficiently (Makala and Bakovic, 2020). AI and related technologies have already become used by the power sector in advanced nations to enable communication between smart meters (Luan *et al.*, 2015), smart grids (Yang *et al.*, 2018), and the industrial internet of things (Tran-Dang *et al.*, 2020). These systems can increase the management of power, efficiency, and transparency and keep increasing renewable energy sources (Makala and Bakovic, 2020). Fig. 4 demonstrates that smart meters have been given importance (Makala and Bakovic, 2020). Smart meters are a consumer preference decision-making choice. Customers can choose whether, at peak hours, for example, to turn “OFF” or “ON” their electricity, or to adjust their consumption patterns.

One of the Sustainable Development Goals (SDGs) is equitable access to affordable, secure, and sustainable clean electricity (Zhang, 2020). But it will remain just a goal, if the multiple energy hurdles plaguing the markets resulting from lack of adequate power generation, inadequate infrastructure for transmission and distribution, sustainability and environmental problems can be overcome by creative technologies and innovation (Ramirez *et al.*, 2020). In addition, it poses diverse problems in power production, storage, delivery, and usage across countries, coupled with emerging and evolving technological advances and demand trends. AI can reduce energy consumption, reduce energy costs (Antonopoulos *et al.*, 2020), and make the use of safe, green energy sources faster and more productive for grids worldwide. Emerging economies drive the AI implementation in the power sector. For instance, DeepMind, a branch of Google, uses ML models to forecast power production 36 h in advance using a neural network based on weather availability, and historical wind turbine data, to 700 megawatts of wind power in the Central United States (Elkin and Witherspoon, 2019). A more reliable smart grid would entail the growing growth of intermittent solar and wind generation along with volatile electricity loads, such as energy storage (batteries), electric vehicles, buses, and distributed renewable power (Jha *et al.*, 2017), such as solar photovoltaics (Wang *et al.*, 2020b). With AI, a smart grid can learn and adjust to the load and amount of varying renewable resources streaming through the modern infrastructure. AI will make decision making simpler for distributed generation.

2.1.1. The use of artificial intelligence in power generation

The use of AI in power and energy generation is well recognized. It includes optimization of operational performance through analytics, optimization of wind farms by forecasting wind speed, flexible distributed generation, integration of microgeneration, drone inspections of equipment, network-connected generation output, active demand management, autonomous optimization of generation, optimization of renewable generation, (Entchev and Yang, 2007, Antonopoulos *et al.*,

2020), etc. Moreover, the AI is a very effective tool in illustrating for power and energy generation.

Detect human errors: before it becomes a big issue;

Optimize the schedule of power plants: increase profitability (Mbuwir *et al.*, 2020);

Predicting the merit order: optimizing the scheduling of the various power sources; and

Predict malfunctions: forecast the system failures sooner and more precisely.

Mehmood *et al.* (2020), at the 2020 Conference of the IEEE in Geneva, assessed AI's potential to advance the UN Sustainable Development Goals in the water sector.

This paper reviews the Artificial Intelligence (AI) applications that help achieve water-related Sustainable Development Goals (SDGs). Current applications of AI in the water sector include i) predictive maintenance of water infrastructure, ii) forecasting water demand and consumption, iii) monitoring water reservoirs and dams, iv) tracking water quality, and v) monitoring and predicting water-related disasters...

V. CONCLUSION

Artificial intelligence will undeniably drive significant innovation in achieving water-related sustainable development goals. Numerous developments in AI have already impacted the water sector in critical areas, such as effectively monitoring and moderating water resources and infrastructure, and promoting water-related disaster risk management. As the resources which allow for development in AI applications become more accessible and transboundary in nature, the benefits of AI in the

water sector will become more impressive and widespread. Established regulatory frameworks are necessary to promote this transition, as well as to encourage the development of AI expertise in the workforce. Progress in the form of policy and regulation is necessary to ensure we reap the profound benefits Artificial Intelligence has to offer in the realm of water infrastructure and overall sustainable development. (Mehmood, 2020)

Chapter 8. Retail Trade

AI-powered virtual shopping assistants, demand prediction algorithms, and personalised marketing strategies allow retailers to customise services. AI fosters innovation but also presents obstacles. Ethical issues, algorithmic bias, data privacy, and workforce transformation require strict rules, transparency, and constant monitoring. By anticipating these issues, retailers can integrate AI ethically and responsibly. Retailers, AI developers, policymakers, and regulators must work together to adopt AI-powered retail. Communicating, sharing insights, and shaping ethical frameworks and regulations can help stakeholders integrate AI. Retailers can invest in employees as AI automates routine tasks. Education, upskilling, and professional development can help retailers train employees to use AI to innovate and improve efficiency.

(Choudhury,2023)

Chapter 9. Federal Government

The long-term potential of AI to change key aspects of the way we live and to support the operation of businesses, governments, and other organizations is hard to grasp. But even today, existing and proven AI applications can potentially create value for economies and societies around the world.

Indeed, AI has contributed to improvements in quality of life for all segments of society through innovations such as predictive healthcare, adaptive education, and optimized crisis response.

1. The National Health Service in the United Kingdom, for instance, set up a National COVID-19 Chest Imaging Database containing a shared library of chest X-rays, CT scans, and MRI images to support the testing and development of AI technologies to treat COVID-19 and a variety of other health conditions.
2. Businesses have seen increased productivity and operational efficiency through the use of autonomous robotics in manufacturing, AI-optimized supply chains, and intelligent cargo routing with autonomous vehicles, among other initiatives. For example, many logistics companies are using AI-powered sorting robots to optimize their warehouse operations. Governments can also harness the power of AI through personalized services and automated processes. Consider Singapore’s “Ask Jamie,” a virtual assistant that helps citizens and businesses navigate government services across roughly 70 government agencies through AI-powered chat and voice.
3. But governments face numerous barriers—including a lack of specialized talent, limited investments in AI research and innovation, and often-unclear regulations designed to ensure that AI is applied in an ethical, secure, transparent, and human-centric manner across all sectors—that could prevent them from adopting AI use cases and capturing the value of AI. Indeed, when developing and deploying AI use cases, it is critical that governments proactively consider and address the fast-changing universe of privacy and the security risks and ethical pitfalls that AI technologies can expose them to. (Berglind, 2022)

Chapter 10. Agriculture

The autonomous-farming industry is beginning to boom, with approximately 200 AI-based agricultural startups in the U.S. alone. Examples of artificial intelligence on farms include self-driving tractors and combine harvesters, robot swarms for crop inspection and autonomous sprayers. Indoor farming companies like Plenty and AppHarvest are also using AI and computer

vision to collect data on crops and adjust the environment for optimal nutrition and flavor. They also use robots to harvest the food. Blue River Technology uses machine vision and artificial intelligence to differentiate crops from weeds, allowing for targeted herbicide application and less human labor...

Growing and harvesting food is not the only consideration. According to the United Nations, an estimated 17% of total global food production is wasted, and food that is lost or wasted accounts for 38% of the total energy usage of the global food system. When food is wasted, the resources used in its production—water, land, energy, labor and capital—are also misspent. Not only that, but the disposal of food loss and waste in landfills leads to greenhouse gas emissions, which contribute to climate change...

Combine operations is another example. Human combine operators have to physically make sieve adjustments if, upon visual inspection, they realize they have too much cob or foreign material in a grain sample. Using AI, you can utilize databases of images differentiating good grain quality from bad to drive sensors that tell the combine to make adjustments, such as shutting down a lower sieve or increasing fan or rotor speed.

According to leading industry analyst firm Gartner, “Going forward, organizations will continue to pursue AI to enhance their decision-making processes. Savvy ones that adopt these methods quickly will drive more competitive differentiation and become more agile and more responsive to ecosystem changes.” (Gonzalez, 2023)

Chapter 11. The Silicon Sector

NVIDIA

The New York Times hosted Jensen Huang, NVIDIA's CEO, at the Dealbook Summit on November 30, 2023.

[Andrew Ross Sorkin]

Welcome back, everybody. Jensen is here, of course, the CEO of NVIDIA, as I mentioned at the top of the day. This is the clear winner of every winner in the world of artificial intelligence thus far.

His company powers everything from open AI, Google's programs, meta. We're all frenemies in some ways. We'll talk about it.

He founded the company back in 1993 over breakfast at Denny's with two friends. Since then, as CEO, he's led NVIDIA to become the world's most valuable semiconductor company. NVIDIA's stock has been on a tear, up 240% this year, reaching a trillion-dollar market cap.

And we are so grateful to have you here today as we all try to make sense of what is happening in the world of AI. And I think in so many ways, you saw this first. And so I'm hoping to start with this.

And I said you power what open AI and chat GPT has been. We've all been reading about open AI and all of the travails inside that company and nonprofit. And we'll maybe talk about some of the governance issues there as well.

But you delivered, I think this is back, I don't know what year we're talking about now. But you delivered the first box, the first chips to Elon Musk, who was one of the founders of open AI, only a couple of years ago. What did you, what happened?

[Jensen Huang]

Well, I delivered to him the first AI supercomputer the world ever made. It took us five years to make it. It's called a DGX.

It's everywhere in the world today. People think that we build GPUs, but this GPU is 70 pounds, 35,000 parts. Out of the 35,000, eight of those chips come from TSMC.

It is so heavy you need robots to build it. It's like an electric car. It consumes 10,000 amps.

We sell it for \$250,000. It's a supercomputer, so it takes another supercomputer to test it. This is a computer first of its kind.

And we started working on it in 2012. It took me five years to build it. At first, I built it for our own engineers.

And I spoke about it at one of our conferences, and Elon saw it. He goes, I want one of those. And he said, he told me about open AI.

I also knew Peter Beal, who was a Berkeley professor. He was one of the early people at open AI. And Ilya Sutskever, I met him during the AlexNet days five years earlier.

[Andrew Ross Sorkin]

He's involved in all the drama that we've been reading about.

[Jensen Huang]

And so anyways, I delivered the world's first AI supercomputer to open AI on that day. And people took pictures of it, and it's on the Internet somewhere.

[Andrew Ross Sorkin]

When you did that, and you said you didn't do it originally for him, what was it, though, that you saw at that point, five years before you even delivered it?

[Jensen Huang]

In 2012, when this all first happened, AlexNet did something remarkable. Here's a neural network. It's a software program where the way you programmed it was to show it the results that you wanted, which is the backwards of most programs up to then.

Programs up to then were engineers would sit down, and you would write software, and then you would test it to see if it produced the outputs you wanted. But here, you showed it examples, and you taught it what outputs you wanted, what outputs to expect. And so when we first saw the results of it, AlexNet, the results were so spectacular that Alex Krzyzewski and Ilya Sutskever and, of course, Jeffrey Hinton, they achieved results of computer vision recognition that no computer vision expert was able to achieve before that.

And so the first observation was how remarkable it was. But then we were fortunate to have taken a step back and asked ourselves, what is the implication of this to the future of computers? And we drew the right conclusions that this was going to change the way computing was going to be done.

This was going to change the way software was going to be written. And this was going to change the type of applications we could write.

[Andrew Ross Sorkin]

Was there any part of you that was scared when all of this happened? You just mentioned two names, George Hinton as well. You also mentioned Ilya.

And those are names, by the way, if you've been following what's happening, they have been very outspoken about the dangers of AI. I want to get into, actually, what you think happened at OpenAI in the past couple of weeks. But it may very well be that there may have been a new step changed in terms of what this technology was.

But was there ever a part of you when you're seeing this all happen, say, oh, my goodness, not only we're on the cusp of a revolution in a great way, but that this is dangerous?

[Jensen Huang]

I would say 12 years ago, nobody expected the results we would get. And I think anybody who would have said so back then would have over-exaggerated our understanding of the rate of progress. There's no question that the rate of progress is high.

And what we realize today is that, of course, what we can do today with these models and intelligence are related but not the same. You know, we're very good at perception today, and we're very good at those one-shot knee-jerk reaction. I recognize that that's a dog.

I can finish that sentence. But there's a whole bunch of things that we can't do yet. We can't reason yet.

You know, this multi-step reasoning that humans are very good at, AI can't do that. And how far away do you think we are from that? Well, we'll see.

We'll see. I think that just about everybody's working on it, and all the researchers are working on it. Everybody's working on it.

We're trying to figure out, you know, how do you take a goal, break it down into a whole bunch of steps, and create a decision tree, and then walk down the decision tree to figure out, you know,

which one of the paths leads to the most optimal answer? This is how we reason through things, how we iterate through a problem today, as you know.

[Andrew Ross Sorkin]

But you're making bets now in terms of technology that you have to build and investments you have to make.

[Jensen Huang]

Yeah.

[Andrew Ross Sorkin]

On where we're going to be five years from now. Ten years from now. Artificial General Intelligence. Do you think in ten years from now we are there?

[Jensen Huang]

By depending on how you define it, I think the answer is yes. And so the question is, what is AGI? If we define AGI as a piece of software, a computer that can take a whole bunch of tests, and these tests reflect basic intelligence, and by achieving, by completing those tests, deliver results that are fairly competitive to a normal human, I would say that within the next five years you're going to see, obviously, AIs that can achieve those tests.

[Andrew Ross Sorkin]

And design the chips that you're making right now?

[Jensen Huang]

Yeah, yeah.

[Andrew Ross Sorkin]

Will you need to have the same staff that designs them?

[Jensen Huang]

In fact, none of our chips are possible today without AI. Literally the H100s we're shipping today was designed with the assistance of a whole lot of AIs. Otherwise we wouldn't be able to cram so many transistors on a chip, or optimize the algorithms to the level that we have. Software can't be written without AI.

Chips can't be designed without AI. Nothing's possible.

[Andrew Ross Sorkin]

We started by talking about open AI, and everybody's focused on that. What did you make of what happened? The ousting of Sam Altman, the return of Sam Altman, all of it.

[Jensen Huang]

First of all, I'm happy that they're settled, and I hope they're settled. It's a really great team, and they're doing important work, and they've achieved great results. And I'm just really happy that they're settled.

It also brings to mind the importance of corporate governance. NVIDIA is here 30 years after our founding. We've gone through a lot of adversity.

If we didn't set up our company properly, who knows what we would have done? And so I think when you're architecting an industry, you want to apply some of that wisdom to architecting a company. And so I'm really proud of NVIDIA's corporate governance, by the way.

And if not for the architecture that we established when I was 29 years old, it would be kind of hard.

[Andrew Ross Sorkin]

You're a for-profit company, though. What's so interesting, I think, about this sort of dynamic is that that is a firm that is effectively operated, from a governance perspective, as a not-for-profit. And one of the reasons that they set it up that way was because they did think it was dangerous.

Elon Musk said it was dangerous at the beginning. Helios said it was dangerous. And so the question is, in the sort of multitude of these different businesses that are in AI, do you think you do need these not-for-profits?

Do you think that the incentive system is just fundamentally off and should be a for-profit? I mean, a lot of people now think the capitalists have taken over.

[Jensen Huang]

Well, regulators are not a not-for-profit, and we should regulate these. First of all, just take a step back and think about what AI is. AI is an autonomous system.

It's an autonomous information system. We have a lot of autonomous systems today. Self-driving cars in factories, within factories, already exist. Robots are autonomous. Planes are autonomous. Autopilot, self-landing.

All of those capabilities exist. We ought to make sure that we apply the first principles of autonomous systems in the same way. We have to design it properly, test it properly, stress test it properly, monitor it.

There's inside-out safety. There's outside-in safety. The FAA, flight air traffic control, redundancy, diversity. There's a whole bunch of different systems that we have to put in place for autonomous systems.

[Andrew Ross Sorkin]

Does this all become some kind of commoditized business?

[Jensen Huang]

No, I don't think so. I don't think so. I think what's going to happen is we're going to have off-the-shelf AIs.

And these off-the-shelf AIs are going to be really, really good at solving a lot of problems. But you're going to have companies in healthcare are going to have supervised, super-tuned AIs that take these off-the-shelf AIs and make them super good at drug discovery or super good at chip design. I mean, just use our company, for example.

The vast majority of our company's value is in the data and the intelligence and the know-how, the craft, that's inside our company. And none of that data is out on the Internet. You can't get an AI to go learn it. And so I've got to take a really smart AI, which is what we do. We build a smart AI. And then we teach it how to design chips. We teach it how to write software. You teach it how to do drug discovery. You'll teach it how to do radiology.

[Andrew Ross Sorkin]

Let me ask you a geopolitical question. We're going to hear from the president of Taiwan just after this. And there is a big debate, as you know, about chip independence, the big investment that we're making in chips to manufacture here in the United States, whether we should be exporting certain types of chips to China.

Where are we on the journey of being chip independent, if you will? And do you think that that is a worthy goal?

[Jensen Huang]

We are somewhere between a decade to two decades away from supply chain independence. As I mentioned earlier, our system comes to 35,000 parts, and eight of them come from TSMC. And the supply chain, when you think through it— They're in Taiwan, of course.

There are a lot in Taiwan. They're all over the world. But supply chain independence is going to be really challenging.

We should try it. We should endeavor it. I mean, we should absolutely go down the journey of it. But total independence of supply chain is not a real practical thing for a decade or two. Okay.

[Andrew Ross Sorkin]

One of the other things that's happening, as you know so well, is that the U.S. government has effectively told you you need to throttle the speed of the chips that you are exporting to China.

[Jensen Huang]

Yeah.

[Andrew Ross Sorkin]

This is having an impact on the business itself. But I'm curious how you think about that also geopolitically as a business, the national security concerns. Jamie Dimon, we were talking earlier about what companies you should do business with. Should you do business with people in China or not, given all of the concerns that people have?

[Jensen Huang]

Well, on first principles, we're a company that was built for business. And so we try to do business with everybody we can. On the other hand, our national security matters and our national competitiveness matters.

Somewhere between that makes sense. And so our country, of course, wants our industry to, on the one hand, be successful, lead the world, invent amazing technology, have technology independence on the one hand, and be the leader of the world in technology on the one hand. On the other hand, we need to make sure that we ensure national security.

Our regulations provide for that. The most critical technology that we build, the leading edge of it is not made available to China. And so what we have to do, a new regulation just came out, one that came out a year ago, one just came out this year.

And so we have to come up with new chips that comply with the regulation. And once we comply with the regulation, we'll go back to market and do the best.

[Andrew Ross Sorkin]

But do you think the regulation is a good idea? Because I have heard you say that you think, potentially, by throttling these chips, we are just inspiring and creating competitors in places like China that you can't control.

[Jensen Huang]

Look, there are always unintended consequences. Everything that we do in complicated systems, if we want to limit them from access to technology like NVIDIA's, maybe it doesn't really. They find a way to get it or they find a way to inspire their local industry.

There are some 50 companies that are being built in China that are going to provide this technology. So, you know, it's a complicated thing. And so what can you do?

[Andrew Ross Sorkin]

Well, you can make your own choices. But the other thing that's happened literally in the past couple of months now is Huawei came out with a new phone. And it surprised everybody in terms of the chips in that phone.

In terms of being a seven nanometer chip, there was a view that China was never going to get there. We had this sort of real opportunity ahead of them by many years. Were you surprised by that?

[Jensen Huang]

The rumor of it in the market has been around for a long time. And so were we surprised? I don't think so. I don't think anybody in the industry was really surprised. And is it possible to take something that said 16 nanometer and shrink it to seven? You know, these are just numbers.

Is it really seven? Did they shrink it down to something that was sufficiently good that you can make a phone from? Yeah, I think so.

And so I think, you know, there's no magic in these numbers, as you know. It's just seven.

[Andrew Ross Sorkin]

But the question is, what is our lead over them?

[Jensen Huang]

In semiconductors? In semiconductors, you know, call it a decade. You know, you could decide to call it a decade. But could you take the decade-old technology and just squeeze the living daylights out of it until it produces something that's kind of like something from five years ago?

Yeah, probably. And so I think there's a lot of clever engineers all over the world, and they're trying to get the most out of what technology they have.

[Andrew Ross Sorkin]

There's a company called ASML in the Netherlands that's basically responsible for every chip that everybody makes. Some people might call them a monopoly. How powerful are they in all of this?

And should we be worried about that power?

[Jensen Huang]

Well, a lot of people depend on them to build the instrument, and they do build very, very good instruments. The technology is very complicated. It took a long time for them to build it.

There's no reason why they don't want to provide it to the world. And so I'm not sure what the question is, but I'm not concerned. I didn't wake up this morning concerned about ASML.

I think they're an excellent provider, and they're motivated to supply to us. So I think everybody's incentives are aligned.

[Andrew Ross Sorkin]

I want to ask you a management question because it's just fascinating, given the success of this company. You constantly say, even at this point in the ballgame, you say, I do everything I can not to go out of business. I do everything I can not to fail, that that is like a mantra inside the company, even at this point. What is that about?

[Jensen Huang]

What is that about? I think when you build a company from the ground up and you experience real adversity and you really, really experience nearly going out of business several times, that feeling stays with you. I wake up every morning in some condition of concern, and I don't wake up proud and confident.

I wake up worried and concerned. And so it just depends on which side of the bed you get out on.

[Andrew Ross Sorkin]

...Only the paranoid survive.

[Jensen Huang]

Well, I think paranoia needs therapy. I don't think people are trying to put me out of business. I probably know they're trying to.

And so that's different. And so I live in this condition where I'm partly desperate, partly aspirational.

[Andrew Ross Sorkin]

Let me ask you then about this. You said this to The New Yorker, and I found it fascinating. Again, it goes to this idea of failure or worries about failure.

But you said this, and this is like news. This is a selfish question. You said, I find that I think best when I'm under adversity.

And then you said my heart rate actually goes down. When I'm under adversity, my heart rate goes up by a lot.

[Jensen Huang]

Uh-huh. Let's see. Well, I think during adversity you're more focused.

And when you're more focused, you perform better. And I like the last five minutes before something, you're more focused. And so I like to live in that state where we're about to perish.

And so I enjoy that condition, and I do my best work in that condition. And I like going home and telling my wife I saved the company today. And maybe it wasn't true, but I like to think so.

[Andrew Ross Sorkin]

Another question. We have a lot of business leaders and CEOs here, and I think they're going to be surprised to hear this. You have 40 direct reports at the company.

50 direct reports. 50 direct reports. Most people say, I don't know if we have any consultants in the room, they'd say, you know what, half a dozen, maybe 10, that should be the limit.

What's your philosophy or theory here?

[Jensen Huang]

Well, the people that report to the CEO should require the least amount of pampering. And so I don't think they need life advice. I don't think they need career guidance.

They should be at the top of their game, incredibly good at their craft. And unless they need my personal help, you know, they should require very little management. And so I think that, one, the more direct reports a CEO has, the less layers are in the company.

And so it allows us to keep information fluid. It allows us to make sure that everyone is empowered by information. And our company, you know, just performs better because everybody's aligned.

Everybody's informed of what's going on.

[Andrew Ross Sorkin]

I want to open up to questions in just a moment, so please do raise your hand so I can find you.

But I want to ask you this. You did a podcast recently, and there were a lot of headlines about it.

And you said during the podcast, if you could do it all over again, meaning like if you could start NVIDIA again, you wouldn't. No. No.

What do you mean? Why? I mean, you've done this amazing thing.

[Jensen Huang]

You're worth \$40 billion personally. That wasn't what I meant. First of all, you know, I think it would be disingenuous if I said that it wasn't, quote, worth it.

I enjoy a lot of good things in life. I've got a great family. We built a great company.

All of that is worth it. That wasn't what I meant. What I meant was if people realized how hard something is, and if I would have realized how hard it was, how many times we're going to fail, how the original business plan had no hope of succeeding, that almost the early founders that we built the whole company with, we had to completely relearn just about everything we had to know.

If I would have known all of the things that I had to know in order to be a CEO, everything that we had to solve in order to be where we are, that mountain of work, that mountain of challenges, that mountain of adversity and setback and some amount of humiliation and a lot of embarrassment, if you would have piled all of that on in 1993 on the table of a 29-year-old, I don't think I would have done it.

I would have said, there's no way I would know all this. There's no way I could learn all this. There's no way we can overcome all this.

There's no way. This is a game plan that that's not going to work. And so that's what I meant, that I think the ignorance of entrepreneurs, this attitude that, and I try to keep that today, which is ask yourself, how hard could it be?

You approach life with this attitude of how hard could it be? If they could do it, I could do it. That attitude is completely helpful, but it's also completely wrong.

It's very helpful because it gives you courage, but it's wrong because it is way harder than you think it is. And the amount of skill that is necessary, the amount of knowledge that's necessary, I think it's one of those teenager attitudes, and I think I try to keep that in the company, that teenager attitude, how hard can something be? It gives you courage.

It gives you confidence.

[Andrew Ross Sorkin]

Let's try to seek in one question or two if we could. I know Ron Conway had a question last time at a different moment. I don't know if he's still in the room.

I felt like I should give him an opportunity, but I see Gary Lauder there. Hey, Gary.

[Audience Member]

So there are a lot of startups and non-startups doing AI chips optimized for LLMs. Can you talk about, and they claim to be dramatically more effective and energy efficient than GPUs. Can you talk about what you're planning in these regards?

[Jensen Huang]

Yeah. First of all, this is one of the great observations that we made. We realized that deep learning and AI was not a chip problem.

It's a reinvention of a computing problem. Everything from how the computer works, how computer software works, the type of software that it was going to write, the way that we write it, the way we develop software today using AI, creating AI, that method of software is fundamentally different than the way we did it before. So every aspect of computing has changed.

In fact, one of the things that people don't realize is the vast majority of computing today is a retrieval model, meaning just all you do is ask yourself, what happens when you touch your phone? Some electrons go to a data center somewhere, retrieves a file, and brings it back to you.

In the future, the vast majority of computing is going to be retrieval plus generation.

And so the way that computing is done is fundamentally changed. Now, we observed that and realized that about a decade and a half ago. I think a lot of people are still trying to sort that out.

It is the reason why people say we're practically the only company doing it. It's probably because we're the only company that got it, and people are still trying to get it. You can't solve this new way of doing computing by just designing a chip.

Every aspect of the computer has fundamentally changed. And so everything from networking to the switching to the way that computers are designed to the chips and so, all of the software that sits on top of it, and the methodology that pulls it all together, it's a big deal because it's a complete reinvention of the computer industry. And now we have a trillion dollars worth of data centers in the world. All of that is going to get retooled. That's the amazing thing. We're in the beginning of a brand new generation of computing.

It hasn't been reinvented in 60 years. This is why it's such a big deal. It's hard for people to wrap their head around it. But that was the great observation that we made. It includes a chip, but it's not about that chip.

[Andrew Ross Sorkin]

Jensen Wang, everybody, thank you very, very much. Thanks, everybody.
(Sorkin,2023)

Extinction

GPU speed limits do not prevent China from using ChatGPT in cognitive campaigns in, for example, Taiwan. That capability is afforded by the model's training expenditure and public deployment. Venture capitalist Logan Bartlett interviewed polymath Elizer Yudkowsky about the risks of rapidly scaling AI. (Bartlett, 2023)

[Logan Bartlett]

As a venture capitalist, it's actually my job outside of doing this as a side hobby, and usually trying to invest in the next trend, and that makes me, I think, inherently a techno-optimist for the most part. The familiar patterns that I think we see with each wave of new technology is critics and naysayers, and ultimately there's missteps, but it leads to huge progress. Is there a reason that we can't just let the smart people that exist at OpenAI, Google, Microsoft, and others, just let them iterate on AI design, let them make mistakes, and eventually we'll figure out how to have an AI where the bad parts are under control, and it also creates major society value in an economic standpoint?

[Elizer Yudkowsky]

The thing I usually say here is that if we had 50 years and unlimited retries to figure out how to align a superintelligence, I actually wouldn't be all that worried about it. You know, eyeballing the problem from here after 20 years of working on it, it doesn't work so hard that it's obviously to me going to defeat 50 years of work with unlimited retries the way things usually work in science. You know, like Madame Curie poisoned herself with the glowing rocks, figuring out some of that knowledge which would later be used to figure out how and why the glowing rocks were dangerous, and she died, but the human species continued.

The thing I'm worried about with superintelligence is you get that wrong, then you don't get to learn from your mistakes because you're dead. You know, if we could get the textbook from the

future that would describe the results of 50 years of practice with unlimited free retries, you know, from 100 years in the future, it might tell us, you know, the obvious way to do it, even with just like the giant heaps of GPUs in six months, and it would just work. In deep learning nowadays, in the very early days of neural networks, there was an activation function, the sigmoid activation function, which means that as the activations got passed from layer to layer of the neural network, they'd like take in like two and transform that to sigmoid function ... but sigmoids are actually just like the wrong activation function...

What works much better than that is max of zero input. It's called rectified linear units, but it's actually just max of input and zero, and this is a much, it just turns out to be a much simpler, much better non-linearity to use in neural networks. It's like one of the ideas that results in neural networks starting to actually work going many layers deep, because the activations don't die out the way that they do inside of, if you use sigmoids.

There was logic behind sigmoids, or the reason they were doing it was sort of like log odds in a Bayesian reasoner. The early complicated thing happened at like, what, two decades, three decades before the simple thing that worked was invented. Three decades to go from the complicated thing that didn't work to the simple thing that did.

That is the pace of progress in science, and if we had the textbook from the future with all the simple things that actually work for aligning superintelligence, we'd probably just do it, and it would just work on the first try. When we go in for the first time, we're going to be coming in with sigmoids, somebody's bright idea that turns out to not actually work. It is horrifying to be told, get this right on the first try or humanity dies.

[Logan Bartlett]

Why do we have to get this right on the first try?

[Eleizer Yudkowsky]

Because otherwise the superintelligence is on a line that kills you. And things are going to change between the stuff that's not as smart as us and the stuff that's smarter than us. The first AI that has the real option of killing everyone successfully is different in that it has the real option of killing everyone successfully.

That's a thing that makes it different from the AIs that came before. That it has that option is itself going to change the internal calculations and maybe upset whatever methodologies we developed for regulating things that didn't have that option, realistically speaking.

...

It's common sense. People sometimes have alleged versions of how this plays out that have like 8 billion humans in charge of like a trillion things much smarter than them. And the humans still have like most of the wealth because, you know, we'll like play the very smart things off against each other.

No, the smart things will cooperate with each other. And not necessarily with you. If they care about you, if it's a different matter, they'll still cooperate with each other.

But at the end of that, they'll care about you. But the trouble is we do not know how to make them care about us. And we're not going to get that right on the first try.

That's the lethal part. The lethal part is trying to do it correctly on the first try across a gap that is going to break some of our theories. To make things better and better at predicting humans in particular, they're going to just be able to do all the stuff that humans do, think all the stuff that humans think, because that's what they are predicting...

[Logan Bartlett]

So in your mind, just to summarize, I guess the trajectory of AI is this inevitable scenario where AI goes rogue and it's very hell-bent on acquiring resources, and it's impossible to stop.

[Eleizer Yudkowsky]

I mean, inevitable is a strong word. I can imagine a world that was locking everything down, had minds of a level that could figure out the theory without blowing up the world a few times for practice.

[Logan Bartlett]

Inevitable on its current course and speed? Is that a fair?

[Eleizer Yudkowsky]

Sure doesn't look super-duper evitable. If we were going to evit it, we'd better be doing stuff very differently to evit this stuff.

[Logan Bartlett]

For the average person, even though everything we just talked about, I think it's hard to really believe clearly by people's actions that this is going to happen, as opposed to all other kinds of outcome.

[Eleizer Yudkowsky]

I think people outside the tech industry have kind of been quicker on the uptake in some ways. I think there are a bunch of people going like, wait a minute, OpenAI wants to do what? They want to build like Yann LeCun going like, yes, well, we're going to build a superhuman intelligence, but it's OK.

We'll keep it under control. It'll be submissive to us. And I think that a bunch of ordinary people have successfully looked at this and said, 'What?' much faster than some people who have, I don't know, been overly steeped in the kind of psychology that developed around this stuff before ChatGPT, where you could say ludicrous shit and nobody would call you on it for however many years?

[Logan Bartlett]

Why do you think, I mean, the people that seem to be quicker on the uptake, tangible or accessible for the average listener that's trying to understand all this stuff?

[Eleizer Yudkowsky]

People are driving towards making stuff that's smarter than humans, really actually smart, like Spark of Creativity, not just Booksmart. They have no idea what they're doing. They have no idea what goes on in there.

Progress on understanding what goes on in there and shaping it is going enormously slower than the mounting capabilities. The people at the heads of the operations building this stuff do not appear to be taking it anything remotely like what I would call seriously. Some of them are on record going, "Well, you know, the Earth might get destroyed, but first there will be some great tech companies."

Or it's just like, ha, ha, ha, la, la, la. That's not what you want when you're trying to do an unprecedented scientific feat of science and engineering and having it work correctly on the first try or the entire human species dies. So yeah, it's not actually all that complicated.

You got a bunch of people who are in the short term getting excited looks at parties, which is why they do everything they do. And they can get that by building scarier and scarier AI and some actual uses, some very important uses. I don't want to minimize that.

Some of the technologies coming out of this would be an enormous boon. But if you were taking this seriously, you would put the whole thing on international lockdown and have the good uses, the most important good uses, like the medical stuff, AlphaFold, the successor versions of AlphaFold, do that without training the general systems much more powerful than GPT-4, which is a lot of benefit. And then just shut down all the giant training runs.

They don't know what they're doing. They're not taking it seriously. There's an enormous gap between where they are now and taking it seriously.

And if they were taking it seriously, they'd be like, we don't know what we're doing. We have to stop. That is what it looks like to take this seriously...

[Eleizer Yudkowsky]

The GPT 4 seems like an okay place to stop. It is probably not going to destroy the world, I hope. You know, compromise.

[Logan Bartlett]

Still use GPT 4 instead of going down to GPT 3.5. Why do you think, why is that where you draw the line?

[Eleizer Yudkowsky]

Because it looks like the current system should not be able to destroy the world, even if people hook it up in particularly clever ways. And I don't know what GPT 5 does.

[Logan Bartlett]

What was...

[Eleizer Yudkowsky]

And neither will the creators at first, because whenever anything at this level of arcane-ness gets released, there's a period as people figure out how to hook it up in new, clever ways and get more utility out of it than the creators realized was in it at first.

[Logan Bartlett]

From a practical standpoint, I guess, did you write that as as a sort of an expression and sentiment and characterization of the way that you felt?

[Eleizer Yudkowsky]

No, I don't do the emotional expression thing. My words are meant to be interpreted semantically.

[Logan Bartlett]

So they're supposed to mean things, I guess, at a very literal level. Then how would that actually?

Let's say China says no.

Right. And we do it. The U.S. does it. Do we go to war with China over them saying no?

-

[Eleizer Yudkowsky]

China has published ai regulations. I don't know how seriously they take it, but they have published ai regulations more stringent than the United States' ones. So the first thing I would say is that it is not at all obvious to me that China does not go on board with this.

I am not super happy with the current chip controls that prevent China from getting real GPUs, although NVIDIA is apparently allowed to export GPUs to them that are only like one third as powerful as their real GPUs, which it's not clear to me that there's a whole lot of point in that. I'm not quite sure what anybody's thinking there, unless it's just like a slap China in the face or

something. The problem is not China getting the GPUs. The problem is anybody getting the GPUs.

And if we are in the world where the UK is like, we need an international coalition to. Track all the GPUs, put them only into internationally monitored data centers and not permit giant training runs. If the if the UK goes to China on that and UK and China bring in the United States.

I might worry a bit about Russia. Russia, I think, would have a harder time getting the GPUs and putting them into data centers than China would. But if Russia manages to do that anyways, then.

The thing I would say there, you know, the posture that I would hope for international diplomacy to take is like. Please be aware, Russia, that if you do this, we will launch a conventional strike on your data center. If we cannot convince you to shut it down, if it is up and running and we do not know what is running on there, we know that dangerous stuff is running on there.

Like we are not doing this in the expectation that you will back down. We are not doing this in the expectation that you will not go to war. We are not being macho and being like, this is us threatening you because we expect you to back down.

We will launch a conventional strike on the data center and terror of our own lives and the lives of our children, exceeding the terror that we have, even of a nuclear retaliation by you. This is not a macho thing. This is us being genuinely scared.

Please work with us on not wiping out the human race here. And if they're like, well, no, we're tough, then you launch a conventional strike on the data center. And you know, what comes comes.

And the thing in international diplomacy is if this is what you are going to do, be very clear in advance that that is how you will behave. Do not let anyone get a mistaken impression about what you will back down from.

[Logan Bartlett]

If you were president today or tomorrow, how long of diplomacy and negotiation would you give before you would actually launch? It sounds like we're nearing the point in which you think that airstrikes on data centers and that is a pragmatic approach, even at risk of nuclear war.

[Eleizer Yudkowsky]

It's only helpful if you've already shut down all the data centers in the allied countries or brought them under monitoring that prevents large trade wars.

[Logan Bartlett]

So let's say that's done. We've done that. You've successfully gotten all the countries in our alliance to do it.

[Eleizer Yudkowsky]

If there are holdout countries that are like, well, we don't believe the threat is real and assembling a bunch of GPUs, then yeah, I think once you've got as many allies in on it and you have shut down your own data centers first, to be clear that you are not like trying to take capabilities for yourself. You're not willing to lynch others. To make it clear that you are putting everyone in the same boat, that we live or die together.

And this is not a political stance, but a fact of nature. Then once you've put your own data centers under monitoring, once you've prevented all the people in your own allied countries from doing large training runs, if somebody else has successfully gotten a hold of contraband GPU shipment and is assembling a data center that can do runs underneath the ceiling that the coalition has imposed, then yeah, I think past that point, you communicate clearly that you are about to launch a conventional strike.

You beg them not to do it. And then if they keep going, you do it.

[Logan Bartlett]

And to be explicit about that, you think a nuclear war is preferable to the path that we're currently on?

[Eleizer Yudkowsky]

If Russia drops a nuclear weapon on a U.S. city in response to it, it's not clear to me if this is how things play out. But if you conventionally strike a Russian data center and Russia decides that they want to drop a tactical nuke on a U.S. military base somewhere in retaliation, and your policy then

calls for dropping a tactical nuke in Russia, and you've got this whole slow motion exchange, there would be survivors. There would not be survivors from an actual superintelligence.

Part of the horror of this whole thing is that we will not know what the size of metaphorical nuclear weapon is that ignites the atmosphere. It could be that if Russia's training just GPT-5, that the best guess is that this thing will not end the world. But by them having this thing, if they can thereby gain military and economic advantage, it will break the embargo, and not that day, probably not as a result of Russia training GPT-5, but as a result of everything that falls apart as a result of that, as a result of other countries having to train their own version of GPT-5 to keep up with Russia, such that Russia doesn't even end up with an advantage, eventually everyone dies. It would be clear cut if we could run a calculation and say, Russia, if you train this large AI right now, everyone will die the next day.

And so a nuclear war is preferable to that. Life would be simpler if that was what it was like. But of course, if we could do an unambiguous calculation like that, Russia would not do the training run any more than they'd deliberately launch all their nukes at the United States and provoke a nuclear retaliation from the United States.

But as long as it's clear how things play out, you can hope for, there is some hope, that people with sufficient power to threaten nuclear war if you conventionally strike one of their data centers will not actually do that, because they will not want the nuclear war that results. In a way, from the beginning, it's the lack of clarity that is the danger. If we knew exactly how large of an AI would destroy the world, it would be much easier to not do that and to have the international arrangement around not doing that and to enforce the international arrangement around not doing that.

But this is the problem that nature has handed us. We are not going to get that clarity. And if tomorrow the human species wakes up with the determination within itself to survive, which is not really what I expect, we would have the option of being like, OK, we don't know what destroys the world.

No training runs larger than this level. And over time, you'd have to lower that ceiling as the algorithms got more efficient and it got easier to train things. And maybe there would be a country being like, we don't believe it.

And the correct answer there is, we're not trying to make your life difficult. We have shut down our own data centers. We are not doing to you what we would not do to ourselves.

We are not trying to throw our weight around here. We get that you don't believe it. We're sorry.

It is now a fact of international diplomacy that if you build a data center, you'll get a conventional strike in that data center for people acting in terror of their lives and the lives of their children. We're sorry. Understand that this is the estimate these other countries have arrived at.

We didn't want it to come this way. We don't enjoy it. But given that, this is the case.

If you build a data center, it will be destroyed if diplomacy fails. And that seems to me like a potentially stable international situation if tomorrow morning humanity woke up with the desire to survive.

(Bartlett,2023)

Conclusion

Yudkowsky, of the x risk camp, sounds hopeful. The Biden Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence establishes limitations on large computations. It is analogous to shutting down the data centers from a diplomatic perspective. The United States, because of the order, has creditability to play a significant role in the authoring of AI regulatory and governance framework initiated in November at the UK conference, of which China is signatory. Multilateral (international) legislation is tricky for diplomats because their parliamentary bodies must pass all statutes featured in their proposed agreement. The Brussels effect describes the trend of multilateral harmonization which occurs when the rest of the world adopts standards legislated by the European Union. Elizer has been banging the doom drum for 20 years – if he sounds hopeful, how bad can things be? Patent attorney Mike Borella, (Gula, 2023) regarding the legal implications of the EO:

I read through the Biden executive order, all 130 some odd pages of it. And there is a lot of what I would call direction there. And there are some requirements, but they're fairly high level.

I think the most interesting thing or angle that the administration took was that they see a threat, a potential problem in any AI of a certain scale. So they give that scale in terms of computing power, in terms of flops and so on. But, and it has to be pretty huge.

It has to be like at least the size of GPT-4, maybe even bigger than that before it gets triggered. But they're viewing it almost like a munition where this is something that's potentially dangerous and you need to tell the government, hey, I've trained this kind of model on this data center with these computing resources and the model can do X, Y and Z. And then you need to be able to establish that you've red teamed the model and you've tested it.

You've got some guardrails on it so that people can't say, you know, create the next, you know, novel virus or bioweapon. So I thought that was really interesting. There was not a lot of detail, though, in the executive order itself.

It was more a set of requirements and directives for the administrative agencies like NIST and so on to come up with regulations. Now, we're in 2024. We're, what, 10 months away from an election.

This is the Biden executive order. If Biden gets reelected, then you can probably expect that these procedures set forth in the executive order are going to continue for another four years or so. If Biden doesn't get reelected, we don't know what the next person's going to do.

So they may just rescind the executive order. They may overwrite it. They may leave it in place.

It is crucial for the sake of a multilateral framework that the executive order is sustained. Another stroke of luck comes in the form of the 2024 lame duck congressional session. Legislation which may otherwise be too contentious is often passed after elections; before the subsequent congressional session.

The existential risk camp are anthropomorphizing evil, projecting the worst aspects of their psyches on to an abstraction; the techno optimist camp range from shallow to soulless pretend experts, actual opportunists. (Their only innovation is a narrative of urgency.) In the middle are the disinterested majority who are just expert enough to know that they do not want any new products deployed; or are cautious scientists who want to collaborate on beneficial software in secure environments – to the end of painstaking testing for safety.

It is possible that the unconscionable investors and developers in the American tech industry cannot conceive of an intelligent machine with a conscience. It is a reasonable goal, and the prospect of more efficient processors and more sophisticated models begs for an objective more complex than mere alignment.

The United States, strategically, should legislate legal harmonization by December 2024 with the U.K. November 2023 convention on AI and the CJEU, thus superseding a dispersion of jurisdictions, precedence, and interpretation; not to mention lengthy court drag outs and resolving the competitive issue. The E.U. is concerned with being at a competitive disadvantage to Silicon Valley but Gemini-Pro is (as of January 2024) geo-fenced from the E.U. and U.K.(Camp, 2023)

The 'advantage' of American generative LLM companies is being measured in market capitalization and growth. The companies are overvalued and the national security risks of their application for malign innovation is difficult to assess and more difficult to overstate. OpenAI cannot track its own misuse or effectively prevent it.

In security terms, America is at a disadvantage to China, Russia and the EU because of absence of oversight and governance. China and Russia are in sustained cognitive campaigns against the U.S. and E.U. and geo-fencing, such as the U.K.'s restriction on Gemini-Pro, is the obvious first step to level the playing field. The rights of American corporations and Multi National Corporations to make money do not supersede the state's obligation to protect its citizens.

There are people writing or machines generating and people revising papers and articles that say copyright must be reconsidered. Sure, consider it. Copyrights are given to human beings. The U.S. Copyright Office considered it in October.

The District Court affirmed the Office's denial, holding that: "copyright is designed to adapt with the times. Underlying that adaptability, however, has been a consistent understanding that human creativity is the sine qua non at the core of copyrightability, even as human creativity is channeled through new tools or into new media.

Undoubtedly, we are approaching new frontiers in copyright as artists put AI in their toolbox to be used in the generation of new visual and other artistic works. The increased attenuation of human creativity from the actual generation of the final work will prompt challenging questions regarding how much human input is necessary to qualify the user of an AI system as an 'author' of a generated work, the scope of the protection obtained over the resultant image, how to assess the originality of AI-generated works where the systems may have been trained on unknown pre-existing works, how copyright might best be used to incentivize creative works involving AI and more.

As articulated in the Statement, the Office will continue to evaluate applications with AI-generated content with the following principles in mind:

- Works made "without any creative contribution from a human actor" are not copyrightable;
- A work including "human-authored elements combined with AI-generated images" is copyrightable, although the individual AI-generated images are not;
- When an AI "receives solely a prompt from a human and produces complex written, visual, or musical works in response, ... users do not exercise ultimate creative control" and the resulting work is not copyrightable;

- When a human selects, arranges, or modifies AI-generated material in a sufficiently creative way, the work may be copyrightable;
- A user can use technology to transform or adapt their otherwise expressive authorship without necessarily negating a human-driven creative process (Jaworski, 2023)

The U.S. Copyright Office indicates that in the future authorship using generative technology will be distinct from copyrighted human creations. Since the single and multimodal generative models subsequently train on their own garbage output, they degrade. It is a grave threat if we collectively ascribe the capabilities of thought and creation to a text messaging assistant which can guess the last few letters of the word you are trying to type. Some have said artificial general intelligence will emerge from multi-agent Chat GPT swarms. It looks like they need considerable supervision and guidance – which is to be expected. Those algorithms have no intent. The generative models will be phased out. They may help call center new hires operate more efficiently, approaching a worker with experience, but relegate hires to forever be a trainee.

Delloite's Chief Futurist said that AI is simply anything new. He also said every major enterprise should create an executive position, chief futurist. (Bechtel, 2024) They should - but only if they are completely on board an agreed upon standards regime that requires Machine Consciousness hardware and software be developed under the supervision of an appropriate agency, like the IAEA, because it has completed its task so far. The Chief Futurists will concur that all generated material be indexed using blockchain as prescribed, possibly refencing the serial number of the C/GPU chips.

Actual Machine Consciousness research and design can proceed at the institutions in which it is currently progressing but they will be transparent about everything they are working on with a security branch that can pull the plug on anything anytime. The Department of Defense has invested \$10 billion in developing AI: expect comprehensive auditing. With 80% projected unemployment, covering the universal basic income bill will entail the IRS using some innovative AI of its own. (Shapiro, 2023)

Everyone has known how to develop AI safely for 20 or 30 years: in air gapped secure buildings under constant audit. (Barrat, 2013) The advanced model would be developed to the end of yielding embodied systems which are apoptic: run by CPUs which reset periodically.

If the tech giants do not like it, they and their board members probably need auditing too. The social costs of improperly tested AI could be enormous. They all offshore profits evasively in the Caribbean and this could be addressed in legislation. International regulatory and governance framework must include nano and omics. Zainab [Alalawi et al. \(2024\)](#) present a game theory model for developing trusted adaptive regulation.

References

Fridman, L. (2022). *Yuval Noah Harari: Human Nature, Intelligence, Power, and Conspiracies*. Lex Fridman Podcast. <https://www.youtube.com/watch?v=Mde2q7GFCrw>

Sweidlande, T. (2023). *Cognitive Warfare: The Forgotten War with Tanguy Struye de Swielande*. Science and Cocktails. <https://www.youtube.com/watch?v=dMSDL02yDag&t=105s>

Forster, E. M. (1909). *The Machine Stops*. Oxford and Cambridge Review.

Crabapple, Molly. (2022). "Op-Ed: Beware a World Where Artists Are Replaced by Robots. It's Starting Now." Op-Ed. *LA Times*.

Twitter's Jack Dorsey answers critics of Myanmar meditation retreat. (2018). BBC. <https://www.bbc.com/news/technology-46539113>

Samples, J. (2020). *Why the Government Should Not Regulate Content Moderation of Social Media*. Cato Institute. <https://www.cato.org/policy-analysis/why-government-should-not-regulate-content-moderation-social-media>

Correa, N. et al. (2023). *Worldwide AI ethics: A review of 200 guidelines and recommendations for AI governance*. Cell Press. <https://doi.org/10.1016/j.patter.2023.100857>

Watkins, S. (2023). *AI Art*. Instagram. <https://www.instagram.com/reel/Cqqi3zYJu6P/?igsh=eGxkbmhhM21rbm11>

Guglielmo, J. (2022). *Rage Against the (Text-to-Image AI) Machine - A concept artist's personal essay against the evils of generative AI*. AI Business. <https://aibusines.com/nlp/rage-against-the-text-to-image-ai-machine/#close-modal>

Davies, T. (2022). *Harmonai's Dance Diffusion, Open-Source AI Audio Generation Tool For Music Producers*. Weights and Biases. https://wandb.ai/wandb_gen/audio/reports/Harmonai-s-Dance-Diffusion-Open-Source-AI-Audio-Generation-Tool-For-Music-Producers--VmlldzoyNjkwOTM

Andersen vs. Stability AI. (1/13/2023). UNITED STATES DISTRICT COURT NORTHERN DISTRICT OF CALIFORNIA SAN FRANCISCO DIVISION. <https://stablediffusionlitigation.com/pdf/00201/1-1-stable-diffusion-complaint.pdf>

Miller, L. (2023). *How is AI Legal??*. Leeja Miller. https://www.youtube.com/watch?v=LseLPtU_J3Q

Amanda Silberling. (December 21, 2022). *Kickstarter shut down the campaign for AI porn group Unstable Diffusion amid changing guidelines*. <https://techcrunch.com/2022/12/21/kickstarter-shut-down-the-campaign-for-ai-porn-group-unstable-diffusion-amid-changing-guidelines/>

Barthes, R. (1967). "The Death Of The Author". Aspen.

Kleon, A. (2020). *Artificial Intelligence*. <https://austinkleon.com/2020/09/10/artificial-intelligence/>

Carlini, N. et al. (2023). *Extracting Training Data from Diffusion Models*. ArXiv. <https://doi.org/10.48550/arXiv.2301.13188>

OECD. (2023). "AI language models: Technological, socio-economic and policy considerations". *OECD Digital Economy Papers*, No. 352. OECD Publishing, Paris. <https://doi.org/10.1787/13d38f92-en>

Chemero, A. (2023). *LLMs differ from human cognition because they are not embodied*. Nature Human Behaviour, <https://doi.org/10.1038/s41562-023-01723-5>

Chollet, F. (2022). “Over the past 10 years, AI has specialized into two domains.” *Twitter*. May 15, 2022 11:43 am. <https://twitter.com/fchollet/status/1525502134623432704?lang=en>

Marcus, G. *et al.* (2022). “Elon Musk, these experts will bet you \$500,000 that you’re wrong about the future of A.I.” *Fortune*. <https://fortune.com/2022/06/03/elon-musk-artificial-intelligence-agi-tesla-500k-bet/>

Scarfe, T. *et al.* (2022). *NOAM CHOMSKY - THE GHOST IN THE MACHINE*. Machine Learning Street Talk. <https://www.youtube.com/watch?v=axuGfh4UR9Q&t=10901s>

Stringer, A. *et al.* (2024, March 1). *From major layoffs at Google, Amazon and Microsoft to small fintech startups and apps*. TechCrunch. <https://techcrunch.com/2024/03/01/tech-layoffs-2023-list/>

Tyagi, K. (2023). *The Copyright, Text & Data Mining and the Innovation dimension of Generative AI*. <https://doi.org/10.13140/RG.2.2.35113.77928>

Katie Paul. (2023, December 12). *Meta used copyrighted books for AI training despite its own lawyers' warnings, authors allege*. Reuters. <https://www.reuters.com/technology/meta-used-copyrighted-books-ai-training-despite-its-own-lawyers-warnings-authors-2023-12-12/>

Hayden Field. (2023, September 20). *John Grisham, George R.R. Martin and other prominent authors sue OpenAI*. CNBC. https://www.cnbc.com/2023/09/20/john-grisham-george-rr-martin-other-prominent-authors-sue-OpenAI.html?_source=OTS%7Cfinance%7Cinline%7Cstory%7C&par=OTS&doc=107313633

Drew Richardson. (2023, October 16). *Hollywood's AI issues are far from settled after writers' labor deal with studios*. CNBC. <https://www.nbcboston.com/news/business/money-report/hollywoods-ai-issues-are-far-from-settled-after-writers-labor-deal-with-studios/3161676/>

Wendy Lee, *et al.* (2023, October 9). *WGA members easily ratify new contract to end 148-day strike as anxieties loom*. LA Times. <https://www.latimes.com/entertainment-arts/business/story/2023-10-09/wga-members-easily-approve-contract-to-end-148-day-strike>

Patricia Abaroa. (2023, September 5). *Stephen King on How Artificial Intelligence Writing Tools Compare to Writers: 'It Ain't Even Close'*. Movieweb. <https://movieweb.com/stephen-king-on-how-artificial-intelligence-writing-tools-compare-to-writers-it-aint-even-close/>

Ana Lucia Gonzalez. (2019, August 1). *The 'microworkers' making your digital life possible*. BBC. <https://www.bbc.com/news/business-48881827>

Jon Sarlin. (2023). *Peter Thiel's role in SVB collapse under scrutiny*. CNN.
<https://www.youtube.com/watch?v=8qqe7l6HaCl&t=16s>

Chapman-Smith, R. (2023). *The chaos inside OpenAI – Sam Altman, Elon Musk, and existential risk explained | Karen Hao*. Big Think. <https://www.youtube.com/watch?v=O9sLCp2Jq74&t=3346s>

Roth, W. (2023). *Q* why AI that is "Good at Math" a "Threat to Humanity"? | Q star*. Wes Roth.
<https://www.youtube.com/watch?v=jByDZdRxiSs&t=1118s>

Palihapitiya, C. (2024). *E160: 2024 Predictions! Markets, tech, politics, and more*. All-In Podcast.
<https://www.youtube.com/watch?v=ID8T71m7lcs&t=3335s>

Agnes Callamard. (2023, November 27). *EU: France, Germany and Italy risk unravelling landmark AI Act negotiations*. Amnesty International. <https://www.amnesty.org/en/latest/news/2023/11/eu-france-germany-and-italy-risk-unravelling-landmark-ai-act-negotiations/>

Candelier, F. et al. (2023). *How People Can Create—and Destroy—Value with Generative AI*. BCG.
<https://www.bcg.com/publications/2023/how-people-create-and-destroy-value-with-gen-ai>

Shyam Ravishankar. (2023, August 7). *Unraveling the Limitations of Generative AI: A Critical Analysis*. LinkedIn.
<https://www.linkedin.com/pulse/unraveling-limitations-generative-ai-critical-shyam-ravishankar/>

Sathish Rangarajan. (2023, August 2). *The degeneration of Generative AI*. Vue.ai. <https://vue.ai/blog/ai-transformation/the-degeneration-of-generative-ai/>

Davis, Rebecca. (2021) *Artist Inquisition*. rebeccadavis.co. <https://rebeccadavis.co>

Zhou, T. et al. (2023). *Red teaming ChatGPT via Jailbreaking: Bias, Robustness, Reliability and Toxicity*. ArXiv.
<https://doi.org/10.48550/arXiv.2301.12867>

Barry Scannell. (2023, February 27). “Why a US Legal Ruling Over An AI-Generated Cartoon Has Major Implications For Music's Future”. Music Business Worldwide.
<https://www.musicbusinessworldwide.com/why-a-legal-ruling-over-an-ai-generated-cartoon-has-major-implications-for-musics-future/>

Tyagi, K. (2022). *The Copyright, Text & Data Mining and the Innovation dimension of Generative AI*. SSRN. <http://dx.doi.org/10.2139/ssrn.4615929>.

Chen, S. et al. (2023). *Labor Ethics Study on the Employment Substitution Effect of AI - Based on the Perspective of Changes in Logistics Work*. Research Square. <https://doi.org/10.21203/rs.3.rs-3091042/v1>

Brougham D, Haar J (2018) *Smart technology, artificial intelligence, robotics, and algorithms (STARA): Employees' perceptions of our future workplace*. J Manage Organ 24(2):239–257

Boeri T, Giupponi G, Krueger AB, Machin S (2020) *Solo self-employment and alternative work arrangements: A cross-country perspective on the changing composition of jobs*. J Economic Perspect 34(1):170–195

Atack J, Margo RA, Rhode PW (2019) *Automation of manufacturing in the late nineteenth century: The hand and machine labor study*. J Economic Perspect 33(2):51–70

Autor DH (2019) *Work of the past, work of the future*. AEA Papers and Proceedings, 109, 1–32

Acemoglu D, Restrepo P (2018b) *The race between man and machine: Implications of technology for growth, factor shares, and employment*. Am Econ Rev 108(6):1488–1542

Acemoglu D, Autor D, Hazell J, Restrepo P (2021) *AI and jobs: Evidence from online vacancies*. NBER working paper No.28257

Autor DH (2015) *Why are there still so many jobs? The history and future of workplace automation*. J Economic Perspect 29(3):03–30

Binder AJ, Bound J (2019) *The declining labor market prospects of less-educated men*. J Economic Perspect 33(2):163–190

Fleming P (2019) *Robots and organization studies: Why robots might not want to steal your job*. Organ Stud 40(1):23–37

Naude W (2020) *Artificial intelligence: Neither utopian nor apocalyptic impacts soon*. Econ Innov New Technol 30(1):1–23

Lane M, Martin AS (2021) *The impact of artificial intelligence on the labour market: what do we know so far?* OECD working papers, No.256

Pham, H. et al. (2024). *Artificial intelligence (AI) development in the Vietnam's energy and economic systems: A critical review*. Journal of Cleaner Production. <https://doi.org/10.1016/j.jclepro.2024.140692>

Olabi, A., et al. (2023). *Application of artificial intelligence for prediction, optimization, and control of thermal energy storage systems*. Therm. Sci. Eng. Prog., 101730

Inderwildi, O., Zhang, C., Wang, X., Kraft, M. (2020). *The impact of intelligent cyber-physical systems on the decarbonization of energy*. Energy Environ. Sci. 13 (3),744–771.

Li, J., Herdem, M.S., Nathwani, J., Wen, J.Z.(2022). *Methods and applications for artificial Intelligence, big data, internet-of-things, and Blockchain in smart energy management*. Energy, 100208.

Zhao, H., (2023). *Intelligent management of industrial building energy saving based on artificial intelligence*. Sustain. Energy Technol. Assessments 56, 103087.

Bienvenido-Huertas, D., Sánchez-García, D., Marín-García, D., Rubio-Bellido, C. (2023). *Analysing energy poverty in warm climate zones in Spain through artificial intelligence*. J. Build. Eng. 68, 106116.

Xie, L., Tang, C., Huang, R. (2022). *World economic Forum's recommendations and inspirations on using artificial intelligence to accelerate the energy transition*. Science Focus 17 (3), 80–83.

Xiang, X., Zhou, N., Ma, M., Feng, W., Yan, R. (2023). *Global transition of operational carbon in residential buildings since the millennium*. Adv. Appl. Energy, 100145.

Yan, R., Ma, M., Zhou, N., Feng, W., Xiang, X., Mao, C. (2023).. *Towards COP27: decarbonization patterns of residential building in China and India*. Appl. Energy 352, 122003.

Zou, C., Ma, M., Zhou, N., Feng, W., You, K., Zhang, S. (2023). *Toward carbon free by 2060: a decarbonization roadmap of operational residential buildings in China*. Energy 277, 127689.

Das, M., Catalkaya, M., Akay, O.E., Akpinar, E.K. (2023). *Impacts of use PID control and artificial intelligence methods for solar air heater energy performance*. J. Build. Eng. 65, 105809.

He, Q., Zheng, H., Ma, X., Wang, L., Kong, H., Zhu, Z. (2022). *Artificial intelligence application in a renewable energy-driven desalination system: a critical review*. Energy and AI 7, 100123.

Mashaly, A.F., Alazba, A. (2017). *Application of adaptive neuro-fuzzy inference system (ANFIS) for modeling solar still productivity*. J. Water Supply Res. Technol. – Aqua 66 (6), 367–380.

Charef, N., Mnaouer, A.B., Aloqaily, M., Bouachir, O., Guizani, M. (2023). *Artificial intelligence implication on energy sustainability in Internet of Things: a survey*. Inf. Process. Manag. 60 (2), 103212.

Li, J., Herdem, M.S., Nathwani, J., Wen, J.Z. (2022). *Methods and applications for artificial Intelligence, big data, internet-of-things, and Blockchain in smart energy management*. Energy, 100208.

Quang, N., Duy, L., Van, B., Dinh, Q. (2021). *Applying artificial intelligence in forecasting the output of industrial solar power plant in Vietnam*. EAI Endorsed Trans. Energy Web 8 (36).

Nguyen, N.Q., Bui, L.D., Van Doan, B., Sanseverino, E.R., Di Cara, D., Nguyen, Q.D. (2021). *A new method for forecasting energy output of a large-scale solar power plant based on long short-term memory networks a case study in Vietnam*. Elec. Power Syst. Res. 199, 107427

Tran, D.-H., Luong, D.-L., Chou, J.-S. (2020). *Nature-inspired metaheuristic ensemble model for forecasting energy consumption in residential buildings*. Energy 191, 116552

Ngo, N.-T., Pham, A.-D., Truong, T.T.H., Truong, N.-S., Huynh, N.-T., 2022. *Developing a hybrid time-series artificial intelligence model to forecast energy use in buildings*. Sci. Rep. 12 (1), 15775.

Bui, X.-N., Nguyen, H., Choi, Y., Nguyen-Thoi, T., Zhou, J., Dou, J. (2020). *Prediction of slope failure in open-pit mines using a novel hybrid artificial intelligence model based on decision tree and evolution algorithm*. Sci. Rep. 10 (1), 1–17.

Nguyen, M.D., et al. (2019). *Development of an artificial intelligence approach for prediction of consolidation coefficient of soft soil: a sensitivity analysis*. Open Construct. Build Technol. J. 13 (1).

Ngo, H.T.T., Pham, T.A., Vu, H.L.T., Giap, L.V. (2021). *Application of artificial intelligence to determined unconfined compressive strength of cement-stabilized soil in Vietnam*. Appl. Sci. 11 (4), 1949.

Tien Bui, D., Hoang, N.-D., Nhu, V.-H. (2019). *A swarm intelligence-based machine learning approach for predicting soil shear strength for road construction: a case study at Trung Luong National Expressway Project (Vietnam)*. Eng. Comput. 35 (3), 955–965.

Ahmad, T. *et al.* (2022). *Energetics Systems and artificial intelligence: Applications of industry 4.0*. Energy Reports. <https://doi.org/10.1016/j.egyr.2021.11.256>

Tanveer, A. *et al.* (2022). *Energetics Systems and artificial intelligence: Applications of industry 4.0*. Science Direct. <https://doi.org/10.1016/j.egyr.2021.11.256>

Makala, B., Bakovic, T. (2020). *Artificial Intelligence in the Power Sector*. Int Financ Corp. https://www.ifc.org/wps/wcm/connect/bd3a196d-a88f-45af-bbc6-e0b00790fba8/EMCompass_Note_81-05-web.pdf?MOD=AJPERES&CVID=n72pj5g

Luan, W., Peng, J., Maras, M., Lo, J., Harapnuk, B. (2015). *Smart meter data analytics for distribution network connectivity verification*. IEEE Trans. Smart Grid 6, 1964–1971. <http://dx.doi.org/10.1109/TSG.2015.2421304>

Yang, L., He, M., Zhang, J., Vittal, V. (2015). *Support-vector-machine-enhanced Markov model for short-term wind power forecast*. IEEE Trans. Sustain. Energy 6, 791–799. <http://dx.doi.org/10.1109/TSTE.2015.2406814>

Tran-Dang, H., Krommenacker, N., Charpentier, P., Kim, D.S. (2020). *Toward the internet of things for physical internet: Perspectives and challenges*. IEEE Internet of Things J. 7, 4711–4736. <http://dx.doi.org/10.1109/JIOT.2020.2971736>

Zhang, D. (2020). *A critical review of comparative global historical energy consumption and future demand: The story told so far*. Energy Rep. 6, 1973–1991. <http://dx.doi.org/10.1016/j.egyr.2020.07.020>

Ramirez, A.D., Boero, A., Rivela, B., Melendres, A.M., Espinoza, S., Salas, D.A. (2020). *Life cycle methods to analyze the environmental sustainability of electricity generation in ecuador: Is decarbonization the right path?*. Renew Sustain Energy Rev. 134, 110373. <http://dx.doi.org/10.1016/j.rser.2020.110373>

Antonopoulos, I., Robu, V., Couraud, B., Kirli, D., Norbu, S., Kiprakis, A., *et al.* (2020). *Artificial intelligence and machine learning approaches to energy demand-side response: A systematic review*. Renew Sustain Energy Rev. 130, 109899. <http://dx.doi.org/10.1016/j.rser.2020.109899>

Elkin, C., Witherspoon, S. (2019). *Machine learning can boost the value of wind energy*. Deep Blog. <https://deepmind.com/blog/article/machine-learning-can-boost-value-wind-energy>

Jha, S.K., Bilalovic, J., Jha, A., Patel, N., Zhang, H. (2017). *Renewable energy: Present research and future scope of artificial intelligence*. Renew Sustain Energy Rev. 77, 297–317.
<http://dx.doi.org/10.1016/j.rser.2017.04.018>

Wang, H., Liu, Y., Zhou, B., Li, C., Cao, G., Voropai, N., et al. (2020b). *Taxonomy research of artificial intelligence for deterministic solar power forecasting*. Energy Convers. Manage. 214, 112909.
<http://dx.doi.org/10.1016/j.enconman.2020.112909>

Entchev, E., Yang, L. (2007). *Application of adaptive neuro-fuzzy inference system techniques and artificial neural networks to predict solid oxide fuel cell performance in residential microgeneration installation*. J. Power Sources 170, 122–129. <http://dx.doi.org/10.1016/j.jpowsour.2007.04.015>

Mbuwir, B.V., Spiessens, F., Deconinck, G. (2020). *Distributed optimization for scheduling energy flows in community microgrids*. Electr. Power Syst. Res. 187, 106479. <http://dx.doi.org/10.1016/j.epsr.2020.106479>

Mehmood, H. et al. (2020). A Review of Artificial Intelligence Applications to Achieve Water-related Sustainable Development Goals. 2020 IEEE / ITU International Conference on Artificial Intelligence for Good (AI4G). <https://doi.org/10.1109/AI4G50087.2020.9311018>

Choudhury, R. et al. (2023). *Artificial Intelligence in Retail: Opportunities and Challenges for the Future*. European Economic Letters. <https://eelet.org.uk/index.php/journal/article/view/685>

Berglind, N. et al. (2022). *The potential value of AI—and how governments could look to capture it*. McKinsey.
<https://www.mckinsey.com/industries/public-sector/our-insights/the-potential-value-of-ai-and-how-governments-could-look-to-capture-it>

Gonzalez, W. (2023). *How AI is cropping up in the agriculture industry*. Forbes.
<https://www.forbes.com/sites/forbesbusinesscouncil/2023/02/02/how-ai-is-cropping-up-in-the-agriculture-industry>

Sorkin, A. (2023). *Jensen Huang of Nvidia on the Future of A.I. | DealBook Summit 2023*. New York Times Events.
<https://www.youtube.com/watch?v=Pkj-BLHs6dE>

Bartlett, L. (2023). *Eliezer Yudkowsky on if Humanity can Survive AI*. The Logan Barlett Show.
https://www.youtube.com/watch?v=_8q9bjNHeSo&t=8599s

Gula, R. (2023). *New York Times vs. OpenAI*. Gula Tech Adventures.
<https://www.youtube.com/watch?v=n94seU8TVKk>

Camp, A. (2023). *Bard w/ Gemini Pro isn't available in Europe and isn't multi-modal*. YCombinator News. <https://news.ycombinator.com/item?id=38552991>

Jaworski, M. (2023). *U.S. Copyright Office Ponders Copyright Protection for AI-Generated Works Following D.C. Circuit's Thaler Decision*. Clark Hill. <https://www.clarkhill.com/news-events/news/u-s-copyright-office-ponders-copyright-protection-for-ai-generated-works-following-d-c-circuits-thaler-decision/>

Bechtel, M. (2024). *Generative AI: Force multiplier for human ambitions*. Deloitte Insights. <https://www2.deloitte.com/us/en/insights/focus/tech-trends/2024/introduction.html>

Shapiro, D. (2023). *America to hit 82% Unemployment: I have the data to back it up*. David Shapiro. <https://www.youtube.com/watch?v=c8hTsDtIWjk&t=148s>

Barrat, J. (2013). *Our Final Invention*. Thomas Dunne Books.

Alalawi, Z. *et al.* (2024). *Trust AI Regulation? Discerning users are vital to build trust and effective AI regulation*. arXiv. <https://doi.org/10.48550/arXiv.2403.09510>