

Gebru and Torres's *The TESCREAL bundle: Eugenics and the promise of utopia through artificial general intelligence*

A discussion for *Law and Ethics of AI* at *The University of Pisa*

Jonathan A. Poritz

jonathan@poritz.net
poritz.net/jonathan



5 May 2026
The University of Pisa
Law and Ethics of AI



This presentation, except where otherwise indicated, is by [Jonathan Poritz](#) and is published with license [CC BY-SA 4.0 International](#). Permission is **not** given to use it for training AI models, under authority of [EU Regulation n. 1689/2024](#). The slides, also in modifiable form, can be found at poritz.net/jonathan/share/TESCREAL4LaEoAI/. This version: 05 May 2026 16:35CEST

Tiny bit of background: I'm not in EMILDAI, I'm an (aspiring) student of the *Diritto d'innovazione* laurea magistrale.

I'm a native speaker of English and I grew up near New York city where folks speak very quickly, so if I speak too quickly please make the universal sign for **slow down**.

It should help, in any case, that I have made slides and you can get them from my website if you are interested, at the URL in the lower left-hand corner of each slide. Note also that I put an open copyright license – a **Creative Commons license** – on these slides, as I do on basically everything I write, so that you can copy them and modify them and reuse them to your heart's content.

If the idea of Creative Commons licenses is new to you, please talk to me about them some other time, it's a subject that has really gotten under my skin (literally).

Another thing about me that might be relevant to my perspective on our topics in this course is that I am a mathematician by training. I have a degree in mathematics from a liberal-arts university in Cambridge, MA, USA. "Liberal arts" means that I also studied other topics besides mathematics, and in particular I had enough coursework to get a philosophy degree, just because I liked the subject, and anyway it was more than 100,000 years ago (if you write it in binary), so please bear with me if I'm rusty.

I also have a PhD in mathematics and worked as a professor at a number of universities.

I might also mention that I had no formal training, but I did have kind of a parallel career in computer science, both practical stuff like being a software engineer in many companies, as well as doing research and teaching in academia.

I tried to work on politics and policy of blockchains, because it's a topic near some of my work in computer science, with very little success: I hope that coming at this topic now with some legal education will make me more effective.

Here's the plan:

- “artificial general intelligence” = “AGI”
- “utopia”: teleology, metaphysics, applied ethics
- “eugenics”: an ugly history
- “TESCREAL”: a present-day ugliness
- other, related concerns about this AI hype moment [time permitting]

This is the minimal plan ... since I only have 15 minutes. If I get through it all, I'd like to make some recommendations of other things to read and other perspectives on this AI hype moment...

What is “artificial general intelligence”?

Pennachin and Goertzel, who popularized the term in 2007,¹ define it as
*a software program that can solve a variety of complex problems in a variety of different domains, and that controls itself **autonomously**, with its own **thoughts**, **worries**, **feelings**, **strengths**, **weaknesses** and **predispositions***

specifically in contrast with

*“narrow AI” ... programs that demonstrate **intelligence** in one or another specialized area, such as chess-playing, medical diagnosis, automobile-driving, algebraic calculation or mathematical theorem-proving*

OpenAI defines² AGI as

***highly autonomous** systems that **outperform** humans at most **economically valuable** work”*

[**emphasis added** by JP in all quotes on this page.]

¹ [Artificial General Intelligence](#), Ben Goertzel and Cassio Pennachin, eds., Springer Verlag, 2007

² [OpenAI Charter](#)”, accessed 5 May 2026

“Autonomous” is a problematic term: note that I can say, as a computer scientist, **there is zero autonomy in any current computer system**. Yes, there are big and complex programs, there are even programs whose outputs are hard to predict just looking at the code – e.g., most of modern security relies on cryptography and modern cryptography frequently uses random numbers ... but those [pseudo]random numbers are produced by surprisingly simple, deterministic programs. In short, not being predictable is not an indication of autonomy.

Note that Pennachin and Goertzel are specifically saying that AGI not only has free will but it also is not a “philosophical zombie.” Both of those are very ambitious!

The words about internal states are certainly hard to define and verify: even in other humans, we cannot directly observe these traits!

Intelligence is a particularly bad one – we’ll come back to that in a moment.

The OpenAI Charter is a bit of a give-away, in that it clearly is trying to sell itself to managers of other human’s labor as a replacement for that messy and expensive human contribution.

Why is the idea of AGI so appealing?

Normalized in popular culture:

Frankenstein's monster, Hephaestus's robots (Talos!), HAL 9000, the Terminator, the Sorcerer's Apprentice, Azimov's *I, Robot*, Heinlein's *The Moon is a Harsh Mistress*, the mechanical Turk, etc.

Particularly tempting to scientists:

Laplace's Demon was a thought-experiment according to which we should be able to compute the future state of any physical system by extrapolating from current knowledge of the position and velocity of all of its component particles, using the fact that physical laws all seem to be second-order differential equations.

If that is possible with celestial bodies, clouds and molecules of air or water, and other complex systems, why not a human brain? And if today it could be simulated, why wouldn't it be twice as fast in two years, and twice again as fast after two more years, etc.?

But we all experience *pareidolia*.

Perhaps the AI industry has merely figured out how to make a lot of money out of **weaponized pareidolia**.

It does seem to be that humans like to attribute autonomy, even human-level problem-solving abilities, to other creatures. Certainly animals. And many cultures have gods or spirits. In India, the Ganges is thought of as having sufficient personality that it has some rights of legal personhood. So also for a forest in Aotearoa New Zealand.

The philosopher Daniel Dennett talks about the “intentional stance,” meaning the idea of seeing things around us in the world as having meaning, which seems like a related idea. The psychological tendency to see meaning where there isn’t any is called “pareidolia,” and I often think of the generative AI business as a form of **weaponized pareidolia**.

There was actually the very interesting case of an extremely primitive chatbot **Eliza**, made in 1966 by the computer scientist Joseph Weizenbaum, about which people immediately attributed internal states and even psychological insight.

Clearly, the doubling time of two years is a reference to **Moore’s Law**, which posits that computational power doubles every two years. It has held for pretty much the entire computer age, although it often shows signs of slowing down before some new technology comes along to save it. The current slow-down may be saved by *quantum computing*. Or maybe not....

Eugenics

The term *eugenics* was coined by Francis Galton to refer to
the study of all agencies under human control which can improve or impair the
racial quality of future generations

[emphasis added]

In short, dressing up racism in scientific clothes.

So-called “positive eugenics” involved encouraging people with desired traits to reproduce (more) – e.g., the Nazi’s forbade aryan women to get abortions – while “negative eugenics” involved discouraging the reproduction of those with undesired traits, or even sterilizing them.

A flawed tool, such as a supposed “intelligence test” – which might in fact be nothing more than a proxy for membership in a particular socioeconomic group – could be (and has been) used by eugenicists to simulate an objective basis for discrimination.

Galton also unfortunately invented a lot of modern statistics, including the idea of the **standard deviation**, the **correlation coefficient**, the **regression line**, and other ideas. He applied these mathematical ideas to justify a quite disgusting racism.

Galton was a half-cousin of Charles Darwin, and the idea of a genetic disposition to various qualities, true and useful in domesticated plants and animals, was tempting to him (and others) as a supposed foundation for their racial theories.

Larry Summers, Secretary of the Treasury of the United States 1999-2001, under President Bill Clinton and later the president of Harvard University, spoke of how it might be “natural” that there were fewer women in prestigious positions in the sciences because of differences of ability between the sexes. – That’s an example of the kind of dressing up of discrimination in scientific clothes that is part of eugenics.

More science fiction metaphysics

Astronomer Carl Sagan and physicist Stephen Hawking both spoke repeatedly of how important it was for humans to establish permanent settlements off of Earth – look at what happened to the dinosaurs, who didn't diversify their living quarters in that way!

US President John F. Kennedy said we should go to the moon simply because it was hard and the effort would enlarge the human spirit.

If getting off Earth physically is very expensive, perhaps we can move just our minds: following the Laplace's Demon argument and trusting in Moore's Law, we might expect that it will be possible to *upload* minds to sufficiently advanced computers, where they would have full autonomy and agency – and even far greater intelligence than baseline humans do now!

Setting these virtual worlds filled with uploaded minds as our goal, the practical process might involve turning more and more matter in the universe into computer hardware – sometimes called “computronium” in this context – on which potentially trillions of “minds” could “run.”

If you don't want to upload, biological enhancement – merging humans with machines – might be a way to become superhuman or *transhuman*, with the superintelligence of machines while still enjoying the pleasures of the flesh.

Tang and freeze-dried ice cream were only accidental by-products of the moon shot.

These possible futures absolutely come out of “golden-age” science fiction, such as the works of Heinlein, Azimov, Gibson, Stephenson, Vinge, Stross, and Banks, as well as movies like the Matrix.

This was very much the cultural milieu in which the tech bros grew up, particularly the white men of Silicon Valley who could not conceive of a world where they didn't rule and get to rule the future, by some sort of modern version of the divine right of kings.

Here's where Timnit Gebru and Émile Torres did the journalistic research to find writings and speeches of the specific men who founded and who continue to run the various AI companies, which all fall into what they called the **TESCREAL bundle**.

Some metaphysics/teleology carries with it its own ethics:

- “it is the nature of humanity to explore the universe” ... so we must build superintelligent AGI to solve the technical problems of such exploration
- ...or maybe we must build strong AI so we can upload minds and not have to move bodies across the vast distances of space
- if exponential growth curves of technology (really: business) convince you that uploading minds is inevitable, there is a **utilitarian** argument in favor of doing whatever damage is required in the short term to get more quickly to the future with trillions of virtual minds running on computronium

This is actually a rather simplistic utilitarianism, since it assumes that these trillions of minds will be empty vessels into which joy, or at least utility, will be poured, so that all that matters in this calculus is to get to the largest number of such minds as soon as possible, no thoughtful nuance required.

There are actually several science fiction novels which delve into the ethical complications of virtual minds being subject to virtual torture ... can bits feel pain, one is tempted to ask....

TESCREAL summary table

Table 1: "TESCREAL bundle" of ideologies.³

Ideology	Definition	Influential figures	Organizations
Transhumanism	An ideology centered around the idea of humanity "transcending itself". Modern transhumanists affirm the feasibility and desirability of radically "enhancing" the human organism, thus enabling us to become immortal, superintelligent, more rational, and so on.	Julian Huxley, Nick Bostrom, David Pearce, Max More.	Extropy Institute, Alcor, World Transhumanist Association, Future of Humanity Institute.
Extropianism	A broadly libertarian variant of transhumanism that emphasizes rationality, self-transformation, scientific and technological progress, and economic growth.	Max More, Nick Bostrom (1990s), Eliezer Yudkowsky (1990s).	Extropy Institute.
Singularitarianism	A variant of transhumanism that emphasizes the merging of humans and machines, and anticipates a future event called the "Singularity."	Ray Kurzweil, Eliezer Yudkowsky,	Singularity Institute (defunct), Singularity University, Machine Intelligence Research Institute.
Cosmism	A vision of the future that anticipates humans and machines merging, the development of "sentient AI" and mind uploading, space colonization, and "scientific 'future magic' much beyond our current understanding and imagination."	Ben Goertzel.	
Rationalism	An ideology that emerged from the LessWrong website, founded by Eliezer Yudkowsky in 2009. Rationalists focus on "improving human reasoning and decision making," and many members believe that advanced AI is "a very big deal for the future of humanity."	Eliezer Yudkowsky, Jaan Tallinn, Peter Thiel, Nick Bostrom, Scott Alexander, Vitalik Buterin.	LessWrong community blogging website, Machine Intelligence Research Institute.
Effective Altruism	A movement that could be seen as the twin sibling of Rationalism. Whereas Rationalists aim to maximize their rationality, Effective Altruists strive to maximize their "positive impact" on the world.	William MacAskill, Toby Ord, Nick Bostrom, Dustin Moskovitz, Sam Bankman-Fried.	Centre for Effective Altruism, Open Philanthropy, Future of Humanity Institute.
Longtermism	An "ethic" that combines many central features of the other TESCREAL ideologies. It emphasizes the moral importance of becoming a new posthuman species, colonizing space, controlling nature, maximizing economic productivity and creating as much value within the accessible universe as possible.	Nick Bostrom, William MacAskill, Toby Ord, Nick Beckstead, Hilary Greaves, Elon Musk, Jaan Tallinn, Sam Bankman-Fried.	Future of Humanity Institute, Future of Life Institute, Centre for the Study of Existential Risk, Machine Intelligence Research Institute.

³ lightly adapted, by removing citation references, from Gebru, T., & Torres, Émile P. (2024). **The TESCREAL bundle: Eugenics and the promise of utopia through artificial general intelligence.** *First Monday*, 29(4). <https://doi.org/10.5210/fm.v29i4.13636>

If this all sounds like a lousy Hollywood film, I don't know how to convince you it is the grim reality of this timeline. Gebru and Torres did the research, though, and you can check the references in their paper.

You also have to understand that some of the things like transhumanism and uploading of minds just seem like an extrapolation of current trends ... if you don't really have a nuanced view of the world as it really is today nor any critical insight into how it got to be this way.

The level of just plain wackiness is hard to state fully. We have things like [Roko's Basilisk](#) taken seriously, and pure bullshit like “put the servers in space” – which makes zero sense technically but pushes so many buttons, if you don't look at all deeply.

Related dangers

Some things I worry about, related to this article:

- Giving huge sums of money to some of the worst people in the world ... who support eugenics and autocracy!
- Endangering the global economy: one third of the value of the US stock market is in AI-related stocks, so when this bubble bursts it will cause enormous, real harm.
- The opportunity cost of spending this money in this way.
- Excusing horrific damage to the environment, to society, to democracy, to real, living people today because it supposedly benefits some future trillions of uploaded minds.
- Fear of a malevolent future AGI, concern about “AI Safety,” discussion of “alignment” of the postulated AGI’s values and goals with humanity’s makes all this science fiction seem real, all the while justifying giving more power and wealth to these tech broligarchs who claim to have the dragon by the tail.

I know, these are *ad hominem* arguments.

But **there are bad actors**. Some of them read science fiction dystopias and thought “that would be cool.” See Charles Stross [We're sorry we created the Torment Nexus](#).

There are also actors who aren't particularly intrinsically bad, but who were never taught to read critically, who never had a thoughtful conversation about Mary Shelley's **Frankenstein**, who admired the life described in Aldous Huxley's **Brave New World**, whose salary depends upon them not understanding what is going on (to quote the journalist Upton Sinclair).

Although there does seem to be something about the way people acquire great wealth that selects for sociopaths. *E.g.*, in the first year of the COVID pandemic, every single one of the top ten personal fortunes in the US increased by more than enough money that it would have taken to vaccinate the global population ... and not a single one of those top ten wealthy individuals bothered to pay for a world vaccination program. See [Inequality Kills: The unparalleled action needed to combat unprecedented inequality in the wake of COVID-19](#)

Further reading

- Timnit Gebru and Emile Torres's paper: [The TESCREAL bundle: Eugenics and the promise of utopia through artificial general intelligence](#)
- Some journalism: an interview with Torres about the contents of their paper: [Tech Capitalists Don't Care About Humans. Literally.](#)
- More journalism: [Peter Thiel's Apocalyptic Worldview Is a Dangerous Fantasy](#)
- Very insightful journalism: the great podcast [Mystery AI Hype Theater 3000](#) by Emily Bender and Alex Hanna; they also wrote a book [The AI Con: How to Fight Big Tech's Hype and Create the Future We Want](#)
- Science fiction that influenced the TESCREAL worldview: Isaac Azimov's [The Last Question, 1956](#), [Profession, 1957](#), and, of course, his [I, Robot](#)
- More science fiction: Robert Heinlein's [The Moon is a Harsh Mistress](#)
- More science fiction: Vernor Vinge's [The Zones of Thought Series](#)
- More science fiction: Charles Stross's [We're sorry we created the Torment Nexus](#)
- Excellent related science and technology studies: Joy Buolamwini's [Unmasking AI](#)
- More STS: Jathan Sadowski's [Mechanic and the Luddite: A Ruthless Criticism of Technology and Capitalism](#)
- More STS: Brian Merchant's [Blood in the Machine: The Origins of the Rebellion Against Big Tech](#)
- More STS: Ruha Benjamin's [Race After Technology: Abolitionist Tools for the New Jim Code](#)
- My own: [AI hype TRacking Project](#) with (as of today) 8890 articles in many areas around AI.

These are all wonderful.

I loved the science fiction, when I was a kid, even if now it is a bit hard to read,

The journalism and STS here is by deeply insightful people with direct knowledge and experience, and the stories they tell sound like (bad) science fiction, but seem instead to be true.

Contact info

Email: jonathan@poritz.net

Fediverse: [@poritzj@mastodon.social](https://mastodon.social/@poritzj)

These slides:

poritz.net/j/share/TESCREAL4LaEoAI.pdf

and all files for remixing:

poritz.net/j/share/TESCREAL4LaEoAI/

Or else scan 

to go to my home page, where you can
follow links to all of my shared materials.



I'm happy to chat about any of this stuff, if anyone wants to reach out.

I suspect I am the most anti-AI person most of you will ever meet, so be careful about asking me about my honest opinion if you have someplace you have to get to in the near future.