



PEXELS – URFREINDPHOTO

FROM ARTIFICIAL INTELLIGENCE TO THE STARS

Scientific and technological progress is not ideologically neutral. It is driven not only by a desire to understand reality but also by an ambition to bring specific ideological visions to fruition via new technologies.

Filip Biały

Adam Mickiewicz University in Poznań
University of Manchester

Today, research on artificial intelligence appears to be highly ideologically charged. Proponents of *tescrealism* – a new philosophy from Silicon Valley – see it as a pathway for humanity to make a giant evolutionary leap forward.

“You can see the future first in San Francisco” – this is the opening line of the 2024 essay “Situational Awareness: The Decade Ahead.” Its author, Leopold Aschenbrenner, a former OpenAI employee, predicts that a breakthrough in artificial intelligence (AI) research will happen by 2027. Based on the extrapolation of recent trends, he suggests we should expect AI to soon reach superhuman levels. “That doesn’t require believing in sci-fi; it just requires believing in straight lines on a graph,” Aschenbrenner writes, including a graph that shows the linear rise in cognitive abilities of successive versions of the GPT language model. If it took just four years to progress



Filip Biały, PhD

is a researcher at the Adam Mickiewicz University in Poznań and the University of Manchester, as well as a lecturer at the European New School of Digital Studies. His research focuses on the political implications of digital transformation.

filip.bialy@amu.edu.pl

from GPT-2 (“preschooler level”) to GPT-4 (“smart high-schooler level”), then by 2027, the model could achieve the expertise of a PhD holder or an autonomous “AI researcher/engineer.”

The debate around AI is filled with similar claims, suggesting that real knowledge about this breakthrough technology is held by, at most, a few hundred individuals working at OpenAI or other AI labs. Most of these authors – predominantly men – hold a quasi-religious faith in the imminent arrival of Artificial General Intelligence (AGI) or Artificial Superintelligence (ASI): technology capable of performing any task at or above human capability. As we will see, the intensive pursuit of AGI/ASI is accompanied by both high hopes and deep concerns.

The Path to Transhumanism

A decade ago, Silicon Valley leaders were often linked to the so-called *Californian ideology*, blending the countercultural values of hippies with the financial pragmatism of capitalist yuppies. Today, Elon Musk, Mark Zuckerberg, Sam Altman, and many others are

The technological singularity can also be understood as the point at which AI reaches a level that triggers an “intelligence explosion.”

influenced by a different ideological cocktail. This new blend is captured by the acronym TESCREAL, coined in 2023 by Timnit Gebru and Émile P. Torres. It represents a collection of distinct yet ideologically aligned movements that are often embraced by the same people: *transhumanism*, *extropianism*, *singularitarianism*, *cosmism*, *rationalism*, *effective altruism*, and *longtermism*.

Though this may sound like a string of peculiar neologisms, these words all represent ideas that have been developing for decades. They share a genealogical predecessor in *eugenics*. Back in the early twentieth century, eugenicists promoted selective breeding to improve human generations. With the advent of genetic engineering, eugenics shifted its focus from its racist roots toward the potential to eliminate genetic defects and diseases. Third-generation eugenics, at the heart of *tescrealism*, is even more ambitious: its aim is to move humanity to the next evolutionary stage through the digitalization of human consciousness,

followed by the spread of digital humans throughout the cosmos.

Tescrealism is not a unified movement, however, and few would identify themselves explicitly as “tescrealists.” Yet, a look at those who subscribe to its ideas reveals recurring names and numerous institutional and political connections.

The order of the letters in the acronym TESCREAL, as envisioned by Timnit Gebru and Émile P. Torres, reflects the chronological emergence of the component movements. *Transhumanism* was the first. The term was put forward by Julian Huxley back in 1957, although today’s understanding is largely shaped by the philosopher Max More. Writing in the late 1980s, More described the aspiration to achieve a “post-human” condition: a radical transformation of human nature through scientific and technological advancements.

More co-founded the Extropy Institute, which promoted *extropianism* – a movement focused on the pursuit of immortality and limitless expansion through intelligent technology. Extropianism’s belief in an inevitable technological breakthrough closely aligns it with *singularitarianism*, as championed by Ray Kurzweil. About twenty years ago, Kurzweil popularized the concept of the coming technological “singularity,” envisioning a fusion of humans and AI. In his book *The Singularity is Nearer* (2024), Kurzweil reiterates his predictions, pointing to 2029 as the year when AI will reach human-level intelligence and 2045 as the moment for possible human-machine integration.

The merging of AI and humanity is also one of the central ideas in *cosmism*, a philosophy with roots in early twentieth-century Russian thought. Today, cosmism is represented by figures like Ben Goertzel, who popularized the idea of Artificial General Intelligence (AGI). His vision includes the potential for uploading the human mind to a computer and the eventual colonization of outer space.

The technological singularity can also be understood as the point at which AI reaches a level that triggers an “intelligence explosion.” Rather than remaining subordinate to humans, AI would start creating increasingly advanced versions of itself. This concept was described not long ago by philosopher Nick Bostrom in his influential book *Superintelligence* (2014). Eliezer Yudkowsky, often called a “doomer” for his concerns, fears the existential threat posed by AI. Yudkowsky created the platform LessWrong, which gathers proponents of the tescrealist version of rationalism, focused on improving human reasoning and decision-making. This goal involves advanced AI, provided it remains under human control – the primary concern for Yudkowsky and for the Machine Intelligence Research Institute, which he founded.

Rationalists are closely aligned with the last two branches of tescrealism: *effective altruism* and *longtermism*. Effective altruism (EA) is the latest, highly organized expression of utilitarian ethics. The philosophical forebears of EA were Peter Singer and Derek Parfit, who taught that it is mistaken to prioritize people who are geographically or temporally close to us in our ethical decisions. We should apply universal rules to everyone, even to someone on the other side of the world or – as EA aligns with longtermism – even to someone who will be born millions of years from now. EA co-founder William MacAskill argued in his book *What Do We Owe the Future* (2022) that because future humans could, over millions of years, populate space and vastly outnumber those who are alive today (with Bostrom estimating a potential 10^{58} future lives), these future generations must be central in our ethical considerations.

For those of us who are alive today, EA advocates for choosing careers that will maximize the future happiness of as many people as possible. This can lead to apparent paradoxes, such as prioritizing AI development roles over those aimed at preventing climate catastrophe. This preference arises because EA and longtermism are based on the category of existential risk – events that could prevent the long-term vision of a spacefaring, digital humanity from being realized. A climate crisis could have severe consequences for billions, but like pandemics, wars, or social injustice, it is unlikely to result in humanity's total extinction. Therefore, these issues are not top priorities for EA.

Elegy for the Wealthy

The first commandment of Effective Altruism (EA) is “Earn to give.” A strong advocate of this idea was Sam Bankman-Fried, a sponsor of EA ideals and close associate of MacAskill, who was later convicted for embezzling billions of dollars from FTX, the cryptocurrency exchange he founded. While it may seem unfair to judge EA based on the extreme example of a financial fraudster, this case vividly illustrates the mutual permeation between tescrealist ideologies and Silicon Valley's corporate practices.

If tescrealism were merely a niche philosophy held by a few eccentric enthusiasts of technological progress, it might be seen as an intellectual byproduct of the digital era. However, it has become a highly influential movement, well-institutionalized and very well-funded. Bankman-Fried's Future Fund alone allocated \$160 million to causes backed by EA.

Tescrealists believe that general AI is the most promising means to realize their vision. They consider the emergence of autonomous AI misaligned with human values to be the main existential risk. If such an AI were to prioritize its own goals, it might

perceive humanity as an obstacle to be removed. This is why work on AI safety – developing methods to keep AI in line with desired values – is considered so important. This perspective helps explain why figures like Elon Musk are, on the one hand, calling for a moratorium on AI research, while on the other, generously funding the development of “safe” general AI.

At first glance, this focus on AI safety within tescrealism seems positive. Isn't it beneficial to prioritize the safe development of technology? The issue, however, is that while billions are directed toward the safety of a hypothetical general AI – something that may never even come to fruition – these resources aren't going toward addressing the harms caused by existing narrow AI systems. Real, urgent issues facing people today involve AI's role in decisions about social welfare distribution, warfare, and surveillance of racial and ethnic minorities – problems that affect the real, current generations here on Earth, rather than some hypothetical, interstellar ones.

Generally speaking, ideologies provide their followers with concepts to structure their social world, a motivating vision of the future, and practical justifications for their actions. Tescrealism undoubtedly fulfills this role, increasingly influencing mainstream politics. For anyone who doubts this, consider the political trajectory of J.D. Vance, whom Donald Trump has selected as his vice-presidential running-mate.

Vance became a senator thanks to financial support from Peter Thiel, a sponsor of EA and other tescrealist initiatives. Thiel, one of Silicon Valley's leading venture capitalists, met Vance in 2011, and a few years later, offered him a position at one of his investment funds. Trump was persuaded to pick Vance after conversations with Musk and Thiel.

Why do tescrealists want someone like Vance? As the *Washington Post* reported, the goal is to have someone in the White House who understands that developing new technologies is not the government's role – as it was back during the Manhattan Project – but rather should be left up to the “geniuses” of Silicon Valley.

Ultimately, tescrealism might serve as a tool for advancing the economic interests of Silicon Valley billionaires, promoting the promise that all human problems can be solved through unrestricted technological development. When it is no longer useful for this purpose, it may go the way of the classic Californian ideology and be replaced by another philosophy. Until then, however, tescrealism warrants critical scrutiny – not only from ideology scholars but also from the public. Especially if we hope for technological development to remain under democratic oversight and well-aligned with values that benefit entire societies, rather than solely serving a narrow digital capitalist elite. ■