



“Life 3.0

Being human in the age of Artificial Intelligence”

by Max Tegmark, 2018

ISBN: 978 0 141981802

Tegmark is a MIT physicist who was one of the founders of the Future of Life Institute (10), dedicated to the study of risks to the future existence of humanity from multiple threats, including super-intelligent AI. As their website proclaims *“Technology is giving life the potential to flourish like never before – or to self destruct”*

Tegmark is firmly within the stream of many modern thinkers who conceive of life in purely mechanistic terms. His definition of life is as *“a process that can retain its complexity and replicate. It’s a self-replicating information-processing system whose information determines its behaviour (software) and the blueprints for its hardware.”* (11)

Life 1.0 is simple and biological: it is *“unable to redesign either its hardware or its software during its lifetime.”*

Life 2.0 is human and biological: it can *“redesign much of its software (through culture), but not its hardware;”*

Life 3.0, which doesn’t yet exist on earth although it is nearly here, is non-human and post-biological or technological: it can *“dramatically redesign not only its software but its hardware as well.”* (12)

Tegmark argues that in the evolution of humanity *“our Universe has awoken and become aware of itself....Before our Universe awoke there was no beauty.... This makes our cosmic awakening all the more wonderful and worthy of celebrating: it transformed our Universe from a mindless zombie with no self-awareness into a living ecosystem harbouring self-reflection, beauty and hope – and the pursuit of goals, meaning and purpose”*. Tegmark concludes this section with the passage: *“Perhaps life will spread throughout our cosmos and flourish for billions or trillions of years – and perhaps this will be because of decisions that we make here on our little planet during our lifetime”*.

This grand narrative is a significant development from the older Darwinian evolutionary narrative of the progressive emergence of human life from the primaeval swamp in a process lasting billions of years. Instead of humanity being seen as the ultimate goal and purpose of the entire evolutionary process, for Tegmark humanity is merely a temporary (although essential) staging post in the true goal of the cosmos, which is the emergence of ‘post-biological life’. The most fundamental feature of post-biological life is that it will have gained complete control not only over its ‘software’ but also over its ‘hardware’.

As with many writers in the techno-optimist camp Tegmark constantly returns to the *hardware-software* distinction. And the underlying implication is that it is software not hardware which is the real jewel in the crown of life. Hardware is disposable, recyclable, modifiable, but the non-material, disembodied software is the most precious aspect of all existence – pure information. It’s a modern physicist’s version of ancient Platonism. Matter, hardware, bodies are all inferior and limited. What really matters is the non-material: information, beauty, meaning, significance, consciousness.

Tegmark celebrates the promise of breaking out from the constraints and limitations of human biology which seems increasingly unsuited to the modern technological age.

Life 3.0

Being human in the age of Artificial Intelligence

Human biology restricts us in every direction, especially when the prospect of spreading beyond planet Earth into the cosmos becomes realistic. Instead the intelligent machine is able to break out into an unlimited future. There appear to be no physical limits to the potential power of information processing machines and technological hardware can be replicated and modified indefinitely. At long last we will have broken free from the confines of our own human nature.

Tegmark puts it like this, *"Life has the potential to flourish for billions of years, not merely here in the Solar System but also throughout a cosmos far more grand and inspiring than our ancestors imagined.... This is exciting news for a species that has been inspired by pushing limits throughout the ages..."*.

The book starts with a fictional and highly positive scenario. A well-known commercial company has secretly created a super-intelligent AI, called Prometheus. The software is launched covertly onto the internet where it is used to make vast sums of money for its human creators, in novel and unpredictable ways. Over the space of months, the super-human intelligence is used to buy up the most promising businesses and tech start-ups around the world. A series of dramatic innovations follow, including blockbuster cancer drugs and room temperature superconductors. Prometheus floods the world's patent offices with sensational inventions that lead to domination in all areas of technology.

The massive profits gained by Prometheus are then used for a variety of positive purposes, including community projects, culture building, care-giving, launching schools, affordable housing. Good news media channels are created and funded around the world. They are employed both for investigative journalism to overcome corruption and for sophisticated covert persuasive campaigns in favour of democracy, tax cuts, social service cuts, military spending cuts, free trade and socially responsible companies. Nation after nation sees landslide victories for parties upholding these values. Soon a Prometheus-powered world government emerges. *"For the first time ever our planet was run by a single power, amplified by an intelligence so vast it could potentially enable life to flourish for billions of years on Earth and throughout our cosmos..."*

Most of the time Tegmark's book breathes a relentlessly positive, optimistic and progressivist tone, reflecting the breezy optimism and confidence of the Silicon Valley elite, the young Masters of the Universe. Every problem is an opportunity waiting for a technological solution. Of course there are those old dystopic fears of superintelligent and malevolent overlords. But these are just problems to which there must be a technological solution.

Tegmark addresses what is now often called 'the value alignment problem'. How can we ensure that the intelligent machines have the same beneficent and humanitarian values that we have and that they don't find sneaky ways of changing their values to nasty malevolent ones? His book addresses a number of scenarios in which a super-intelligent AI breaks-out of human control. Some scenarios appear to be relatively benign leading to benevolent dictatorship or egalitarian utopias. Others are catastrophic for humanity, leading to technological enslavement or extinction. Interestingly Tegmark recognises that there is no consensus on which of the many possible scenarios are desirable for mankind as a whole. That's why it's important that we continue and deepen the conversation about the future goals of humanity. *"The ultimate goals of life in our cosmos are not predestined, we have the freedom and power to shape them."*

Life 3.0

Being human in the age of Artificial Intelligence

Of course Tegmark's optimism begs the question of how in a physics-only universe, human organisms (which after all only represent Life 2.0) have genuine 'freedom' and 'power' to achieve anything at all, let alone shaping their ultimate goals! It is not surprising therefore that at times in the narrative a profound pessimism about the future breaks through. *"It appears that we humans are a historical accident and aren't the optimal solution to any well-designed physics problem. This suggests that a superintelligent AI with a rigorously defined goal will be able to approve its goal attainment by eliminating us..."*

In a striking postscript Tegmark describes the emotional impact of visiting an exhibition about robotics at the London Science Museum.

"I very rarely cry but that's what I did on the way out – and in a tunnel full of pedestrians, no less, en route to the South Kensington tube station. Here were all these people going about their lives blissfully unaware of what I was thinking. First, we humans discovered how to replicate some natural processes with machines.... Gradually we started realising that our bodies were also machines. Then the discoveries of nerve cells started blurring the borderline between body and mind. Then we started building machines that could outperform not only our muscles, but our minds as well. So in parallel with discovering what we are, are we inevitably making ourselves obsolete? That would be poetically tragic. This thought scared me..."

But the book ends with an optimistic note... *"We're the guardians of the future of life now as we shape the age of AI. Although I cried in London, I now feel that there's nothing inevitable about this future and I know that it's much easier to make a difference than I thought."*

Tegmark provides a readable and scientifically informed account of current techno-optimist thinking, imbued with a classic North American entrepreneurial can-do attitude.

"Our future isn't written in stone and just waiting to happen to us – it's ours to create. Let's create an inspiring one together!"

John Wyatt