5

The World as the Perfect Machine Universe

Contemplate the whole and every part of it: You will find it to be nothing but one great machine, subdivided into an infinite number of lesser machines.
David Hume

When genetic designer J.F. Sebastian from the movie *Blade Runner* (Ridley Scott. USA, 1982) discovers Pris, a humanoid robot woman on a rainy night in a pile of garbage, he invites her to his home. Sebastian lives in a gloomy run-down building sparsely illuminated by floodlights.

"It must get lonely here," Pris says. But Sebastian denies that. He makes his own friends, he says. When the two of them enter his apartment, a little electric soldier with a long red nose and a little animated teddy bear come to greet him.

"Welcome home," they say. Now, the spectator understands what he meant by his comment. His "best friends" have indeed been made by him. They are mechanical dolls, uncanny automatons, kind of alive but also dead. In his essay "Animism, Magic and the Omnipotence of Thought" (1913), Freud describes the belief in the ensoulment of plants and artificial as well as natural objects as an "animistic system of thought" based on magical ideas. Behind this, according to Freud, lies both the early childhood desire for omnipotence and the refusal to grow up. Those who believe in animate objects regress into childish fantasies of omnipotence and suffer from a narcissistic disorder**.**

Automata have already fascinated people in antiquity. The automata at that time were created on the basis of hydraulics and pneumatics, such as a small arrow-shooting Heracles by Heron of Alexandria or a life-size automaton by

J. Nida-Rümelin, N. Weidenfeld, *Digital Humanism*,
https://doi.org/10.1007/978-3-031-12482-2_5

Philon of Byzantium pouring wine and water as early as the second century BC. This fascination with life-like, mechanical creatures remained in the course of later centuries. Particularly in the seventeenth and eighteenth centuries, where the manufacture of automata became increasingly precise and impressive thanks to the developments in the art of watchmaking.

Behind this fascination lies more than just a superficial enthusiasm for mechanics. It is the idea of the world as a clock. In a rejection of traditional authorities and of the traditional Aristotelian-Thomistic Catholic worldview, the hope for a complete explicability and formability of the world increasingly develops from the sixteenth century onwards. The powerful movement of *Scientia Nova* emerged: revolutionary thinkers dedicated solely to the scientific-rational argument. This rationalism is modelled on the mathematical method of Euclid (*more geometrico* = in the manner of the geometric mathematician as the ideal of philosophy). Leibniz, the brilliant philosopher and mathematician of the pre-Kantian period, developed the idea of a universal calculating machine and understood the rationally ordered world as an expression of divine creative will. With the help of logical conclusions and mathematical methods, it should be possible to calculate every event in the world. The world as a whole is understood as a deterministic system according to strict mathematically describable laws.

It is our bold but not entirely far-fetched conjecture that we are, at present, entering a new era of rationalism that now expects from Artificial Intelligence what the rationalists of the seventeenth century lacked, namely, the means for a complete rational collection and processing of all data. The hope is that one day our entire living world will be permeated by technological-scientific rationality: every area illuminated, rationally ascertainable, and predictable. We suspect this to be an expression of an unconscious wish to counterbalance the fact that life is often enough chaotic and unprogrammable, but this would be up to psychanalytic theory to investigate further. The image of the self-thinking robot remains a kind of emblem for this (old) rationalist hope.

Rationalists make no distinction between artificial and human intelligence. They stand for a position called "strong AI."[1]

Strong AI implies the thesis that there is no (categorical) difference between human thought and software or computer processes (computing). These two types of thought processes not only follow the same rules but do not differ in any essential respect, so that it makes no sense to reserve the mental vocabulary (notions like beliefs, desires, perceptions, feelings, etc.) for only one of the

[1] An interesting proponent of strong AI is Milkowski (2013).

two types. The simplest interpretation is behavioristic: being sad means nothing more than exhibiting a behavior characterized as sad.

The proponents of a so-called strong AI consciously or unconsciously advocate the ideal of the universal, completely determined machine as an explanatory pattern of the world and of humankind. Strong AI in all its variants is a form of anti-humanism. It negates both human reason, i.e., the ability to be guided by reasons, and the role of subjective mental states in a part of animate nature. Strong AI is logically incompatible with both the existence of *qualia* (*qualia* are states of feeling, such as what it is like to perceive something, for example, the color red) and the existence of objective reasons. Strong AI is the contemporary variant of a crude, mechanistic materialism. Such a materialism degrades the human individual to a digital, mechanical system that can be determined and predicted by sensory stimuli and thus falls behind the achievements of humanism.

Exactly such a nightmare is imagined in *The Matrix* (Lilly and Lana Wachowski. USA, 1999), where machines have taken over and keep the humans as predictable energy sources. In the final showdown in the third part of *The Matrix* trilogy the hero of the film, Neo, enters the machine world and faces the all-powerful master of the machines, a kind of mechanical kind of god made out of millions of small mechanical parts. This "god," who speaks in a deep electronic voice, has no empathy whatsoever: neither for Neo nor for the rest of humanity. His goal is to keep the machine world functioning smoothly. Now that humans have begun to develop a will of their own, he would rather like to get rid of them.

This machine god is a perfect symbol of the ideology of the world as a machine and what the film tells us at this point is that a world run by such an ideology can only lead to an inhumane world.

In addition to the strong AI position, there is so-called weak AI position. This can also be found in AI discourses. This position does not deny that there are categorical differences between human and Artificial Intelligence but claims that there is no fundamental limit to the computerization (digitization) of human thought, perception, decision-making, and feeling. Weak AI assumes that in principle all human thinking, perception, and decision-making processes can be *simulated* by suitable software systems. From a humanistic point of view, weak AI is therefore ruled out as an alternative to strong AI, because how can the differences between human and Artificial Intelligence be determined at all if all human abilities can in principle be simulated? As a counter-model to the anti-humanist strong AI, weak AI is just that: too weak. The only plausible alternative to the strong AI ideology and its implicit mechanistic thinking is digital humanism. A humanism that neither doubts nor

threatens human authorship, but rather expands it through the use of digital technologies.

The boom in neuroscience has given new impetus to an anti-humanistic mechanistic worldview. When, for example, they use computer tomography to visualize which part of the brain is being supplied with blood when someone decides to drink a cup of coffee, they conclude that it is the brain, or rather neurophysiological states, and not the person as an agent, that determines the action. But this is a fallacy: showing that actions or intentions are accompanied by patterns of blood flow and activation in specific brain regions does not mean that our actions are caused by these physiological states, nor does it mean that we really understand *how* this processing takes place. The observation of a neuronal correlation must not lead us to the (mechanistic) ideology that all human decisions can be identified with brain activities.

The operation of reasons is central to the (humanist) human self-understanding.[2] Humanists are fallibilists, that is, they consider it possible that any of our beliefs could also turn out to be false under certain conditions. We do not invent our world through deliberation, but we try to approach it in this way in order to understand it better.

In a humanistic worldview, a human being is not a mechanism, but a free (autonomous) and responsible agent in interaction with other human beings and a shared social and natural world. He is not merely part of a great machinery, a cog in a wheel, not an optimizing monad moved by sensory stimuli, but self-effective in a world moved only in part by mechanical relations. Analogous to the medieval conception of God as an unmoved mover, man is an agent. A multiplicity of unmoved movers, of persons who intervene in and shape world events according to their own evaluative judgments, constitutes a humane society.

At the beginning of the twentieth century, there was another science fiction film that, like *The Matrix*, focused on the inhumanity of a world ruled by a machine, or metaphorically speaking a world ruled by the ideology of a mechanistic worldview: Fritz Lang's film *Metropolis* (Germany, 1927). In the world of *Metropolis,* the wealth of a few who live in a luxurious upper world is acquired by the work of many who live in the underground working with machines, who produce goods and energy. The inhumanity of their work lies in the fact that these workmen are degraded to robots themselves through their work, as they are required just to function and work, without communicating with others and without creating social bonds. The human workers thus function according to the beat given by a super-machine, staged

[2] For the philosophical underpinning of this account, see Nida-Rümelin (2023).

by Fritz Lang as a kind of cruel machine god, who demands absolute devotion from the human workers—even if it means their complete exhaustion and often their death. For this machine god, only efficiency and performance count. Human lives have no meaning.

At the end of the film after a major confrontation between the two worlds has occurred, during which the underground world of the workers is destroyed, the workers, who have lived in caves and underground cities all their lives, step out and come up the surface for the first time meeting their masters. With the help of the protagonist, who wants to bring the two worlds together, a new way of cooperation and a new beginning seems possible.

In the *Allegory of the Cave*, Plato tells us about people who spend their lives in a cave. Much like the workmen of *Metropolis* they have never seen the sun and the world above the cave and thus do not know what the world is truly like. All they see are shadow images of things thrown up on the wall by the light of a fire behind them.

If we cling to a mechanistic view of the world, we deprive ourselves—just like Plato's cavemen—of the possibility of taking a true look at the world, which is much more than just a small wheel in the gears of a great universal machine.